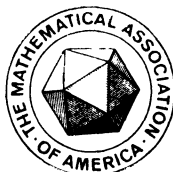


THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF
THE MATHEMATICAL ASSOCIATION OF AMERICA



VOLUME 80

NUMBER 1

CODEN: AMMYAE

CONTENTS

Unique Factorization Domains	P. M. COHN	1
Continuous Analogues of Series	R. P. BOAS, JR. AND H. POLLARD	18
England was lost on the Playing Fields of Eton: A Parable for Mathematics	A. B. WILLCOX	25

MATHEMATICAL NOTES

An Identity Satisfied by Derivations of a Purely Inseparable Field	F. P. CALLAHAN	40
On Sums of Powers of a Number	VLADIMIR DROBOT	42
A Local Mean Value Theorem for Analytic Functions	ÅKE SAMUELSSON	45
A Theorem on Set Inclusion in Metric Spaces	J. A. HEINEN AND ALBERT WILANSKY	46
Circle Groups of Nilpotent Rings	J. C. AULT AND J. F. WATTERS	48

RESEARCH PROBLEMS

Crossing Number Problems	P. ERDÖS AND R. K. GUY	52
------------------------------------	------------------------	----

CLASSROOM NOTES

A Proof of Uniqueness of Factorization in the Gaussian Integers.	M. F. RUCHTE AND R. W. RYDEN	58
Some Half-plane Dirichlet Problems: A Bare Hands Approach	F. J. FLANIGAN	59
Single Layer Potentials and the Cauchy-Kowalewski Theorem	P. A. NICKEL	61
A Global Characterization of Uniform Continuity	RICHARD CLEVELAND	64

MATHEMATICAL EDUCATION

Applied Mathematics at M.I.T.	H. P. GREENSPAN	67
A Letter by Professor Polya		73

ELEMENTARY PROBLEMS AND SOLUTIONS		74
ADVANCED PROBLEMS AND SOLUTIONS		82

(Continued on inside cover)

JANUARY

1973

REVIEWS	88
NEWS AND NOTICES	114
MATHEMATICAL ASSOCIATION OF AMERICA	114
May Meeting of the Indiana Section	114
Employment Information for Mathematicians	115
Calendars of Future Meetings	116

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss.

Backlog: Main Articles 12 months, Math. Notes 15 months, Research Problems 7 months, Classroom Notes 11 months, Math. Education 10 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to HARLEY FLANDERS, American Mathematical Monthly, Tel Aviv University, Ramat Aviv, Israel (see Notice, vol. 77, 1970, p. 555); NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D. C. 20036.

HARLEY FLANDERS, *Editor*

ASSOCIATE EDITORS

JOSHUA BARLAZ	J. G. HARVEY	SEYMOUR SCHUSTER
E. R. BERLEKAMP	ERIC S. LANGFORD	J. ARTHUR SEEBACH, JR.
JANE W. DI PAOLA	P. D. LAX	E. P. STARKE
ROBERT GILMER	ARTHUR MATTUCK	LYNN A. STEEN
RICHARD GUY	M. W. POWNALL	JAMES WENDEL
RAOUL HAILPERN	GIAN-CARLO ROTA	

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are \$12.50. For nonmembers the subscription price is \$18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices.

Copyright © The Mathematical Association of America (Incorporated), 1973

PRINTED IN THE UNITED STATES OF AMERICA

UNIQUE FACTORIZATION DOMAINS

P. M. COHN, Bedford College, University of London

1. Introduction. One of the most fascinating developments in ring theory in recent years is the way in which large parts of algebraic geometry can now be stated entirely in terms of commutative Noetherian rings [23]. In the other direction this has led to new tools for classifying and investigating these rings; furthermore, these applications are no longer confined to Noetherian rings, and although commutativity is assumed as a rule, one suspects that even this is not always essential. There is an extensive and rapidly growing literature on the subject, and it would be difficult to do justice to it in a brief article. Instead, we have singled out a special class of rings: unique factorization domains. They provide a good example of how ring-theoretical properties can be illustrated by geometrical ideas. The non-commutative case is described separately; this is less well developed, and the connection with geometry is less clear, but eventually any geometric ring theory must also comprehend the non-commutative case.

We begin with the definition of a commutative unique factorization domain (UFD) in section 2, and its relation to such basic notions as Dedekind and Krull domains. In section 3 we analyse the definition and reduce it to a statement about primes. Nagata's theorem and its application to Felix Klein's theorem on line complexes is discussed in section 4.

The remainder of the article is concerned with the non-commutative case. In section 5 we discuss the lattice-method of defining and recognizing UFD's and give some non-commutative examples. All are consequences of the fact that an atomic 2-fir is a UFD. If we drop atomicity, we are left with the Schreier refinement property, which so far has mainly been studied in the commutative case (section 6). Section 7 describes two special cases of interest in the non-commutative theory: the lattice of factors is distributive, respectively a chain. In section 8 we examine the shortcomings of the definition of non-commutative UFD and describe some remedies that have been proposed, and section 9 notes the problems of factorizing zero-divisors.

Throughout, some of the easier proofs have been sketched and others omitted, with a reference where full proofs can be found. When no convenient reference was available, proofs are given in more detail. The article is based on a lecture delivered at the British Mathematical Colloquium at Leicester on April 1, 1964.

Professor Cohn did his Cambridge Ph. D. under Philip Hall, and he has held positions at the Univ. de Nancy, Manchester University, Queen Mary College London, and (presently) Bedford College, London. He has spent leaves-of-absence at Yale Univ., Univ. of Chicago, and Rutgers. He has published extensively in universal algebra and in many branches of algebra; in 1965-67 he was the Secretary of the London Mathematical Society. His Books are *Lie Groups* (Cambridge Univ. Press 1957), *Linear Equations* (Routledge and Kegan Paul 1958), *Solid Geometry* (Routledge and Kegan Paul 1961), *Universal Algebra* (Harper and Row 1965), and *Free Rings and Their Relations*, LMS Monograph 2 (Academic Press 1971). He received a MAA Lester R. Ford Award in 1972.

Editor.

2. Commutative unique factorization domains. All rings are understood to be associative, with a unit-element 1, which is inherited by subrings and preserved by homomorphisms. Moreover, all modules are unital. Usually our ring R will be an **integral domain**, i.e., the set R^* of non-zero elements is non-empty and closed under multiplication. This terminology will be used even for non-commutative rings (though at first our rings will be commutative). An element u of a ring R is called a **unit**, if there exists $v \in R$ such that $uv = vu = 1$; any non-unit which cannot be written as a product of two non-units is said to be **irreducible** or an **atom**. An integral domain is said to be **atomic** if every element, not zero or a unit, is a product of atoms.

DEFINITION. A commutative integral domain R is said to be **factorial** or a **unique factorization domain (UFD)** if it satisfies the following conditions:

A. R is atomic,

U. Any two factorizations of an element into atoms differ only in the order of the factors, and by unit factors.

Thus if $c = a_1 \cdots a_r = b_1 \cdots b_s$, where a_i, b_j are atoms, then $r = s$ and after a suitable renumbering of the b 's, a_i is associated to b_i , i.e., $a_i = b_i u_i$, where u_i is a unit.

The best-known example of a UFD is the ring \mathbf{Z} of integers. There are two basic ways of proving that \mathbf{Z} is a UFD, which we shall call the **prime method** and the **lattice method**. Both are capable of generalization and we shall deal with each in turn (sections 3 and 5).

UFD's are important for several reasons: In the first place, their characteristic property makes them more amenable; secondly, to impose unique factorization often singles out a significant class of rings, while thirdly, the methods used to prove factoriality have often given rise to other notions important in their own right.

The unique factorization property of the integers can also be shown to hold for the ring of Gaussian integers $a + b\sqrt{-1}$ ($a, b \in \mathbf{Z}$), and this led to efforts to prove the same for the ring of integers in any finite algebraic extension of \mathbf{Z} . These efforts, though doomed to failure, led Kummer to introduce 'ideal numbers' in an attempt to restore unique factorization. Dedekind [18] gave a general definition of ideals and showed that rings of algebraic integers possess unique factorization for their ideals. Thus if \mathfrak{p}_i ($i \in I$) are the different prime ideals, any non-zero ideal \mathfrak{a} has a unique representation

$$(1) \quad \mathfrak{a} = \prod \mathfrak{p}_i^{v_i},$$

where the integers $v_i = v_i(\mathfrak{a})$ are non-negative and all but a finite number of them are zero. A commutative integral domain with unique factorization of ideals is called a *Dedekind domain*; such a ring is necessarily Noetherian, i.e., it satisfies the ascending chain condition, briefly ACC, for ideals. Any Noetherian UFD is a Dedekind domain, but there are UFD's that are not Noetherian, and hence not Dedekind, e.g., the polynomial ring in infinitely many indeterminates over a field.

To get a common generalization of UFD and Dedekind domain, we observe that in both cases we have a family of integer-valued functions on R^* satisfying the familiar conditions for an exponential valuation:

$$V.1. \ v(a) \geq 0,$$

$$V.2. \ v(a - b) \geq \min\{v(a), v(b)\},$$

$$V.3. \ v(ab) = v(a) + v(b).$$

Such a valuation extends in a unique way to the field of fractions K of R . Now Krull [30] considered more generally, integral domains with a family $(v_i)_{i \in I}$ of such valuations, where for any $c \in K^*$, $v_i(c) = 0$ for almost all i , and $c \in R$ if and only if $v_i(c) \geq 0$. Such rings are called *Krull domains* (Krull called them "endliche diskrete Hauptordnungen"); clearly they include both UFD's and Dedekind domains as special cases, e.g., a Noetherian integral domain is a Krull domain if and only if it is integrally closed (in its field of fractions). Krull domains retain at least some of the useful features of UFD's; moreover, unlike UFD's, the class of Krull domains is closed under integrally closed integral algebraic extensions [6]. For any Krull domain its departure from factoriality is measured by the *divisor class group*, i.e. the group of all divisors (formal products $\prod p_i^{v_i}$) modulo the principal divisors [36]. Its vanishing characterizes UFD's; it is unchanged under adjunction of indeterminates, but may change under algebraic extension.

The relation between the different types of ring becomes clearer if we adopt a slightly different point of view. Let K be any commutative integral domain and K its field of fractions. On K^* we define the relation of divisibility: a divides b , in symbols: $a \mid b$, if $ba^{-1} \in R$. Clearly this relation is reflexive and transitive, i.e., it is a *preordering* of K^* . Moreover, it is compatible with multiplication: if $a \mid b$, then $ac \mid bc$ for all $c \in K^*$. In this way K^* becomes a preordered group; if U is the group of units in R , then $D = K^*/U$ is the partially ordered group associated with the preordered group K^* ; it is called the *divisibility group* of R . Now various classes of rings can be described entirely in terms of the order type of their divisibility group. For comparison we shall need ${}^I\mathbf{Z}$, the direct sum of $\text{card}(I)$ copies of \mathbf{Z} , with the componentwise ordering: $(x_i) \leq (y_i)$ if and only if $x_i \leq y_i$ for all $i \in I$. Then

- (i) R is a UFD if and only if D is order-isomorphic to ${}^I\mathbf{Z}$, for some I ,
- (ii) if R is a Krull domain then D is order-isomorphic to a subgroup of ${}^I\mathbf{Z}$, for some I ,
- (iii) R is a valuation ring if and only if D is totally ordered,
- (iv) R is a discrete valuation ring if and only if $D \cong \mathbf{Z}$.

3. The relation of UFD's to primes. Let us analyse the notion of UFD more closely. An element p of a commutative integral domain R is said to be *prime* if p is not zero or a unit and $p \mid ab$ implies $p \mid a$ or $p \mid b$. From this definition it is easy to see that each prime must be an atom. The converse need not hold; in fact with a finiteness condition it is equivalent to unique factorization:

THEOREM 1. *A commutative ring is a UFD if and only if it is an atomic integral domain and every atom is prime.*

This is a sort of localization of the condition U . It is easily proved; in fact the usual proof that \mathbf{Z} is a UFD consists in verifying that every atom is prime, using the Euclidean algorithm, and then carrying out what is in effect a proof of Theorem 1 [42].

Examples of atomic integral domains that are not UFD's are well known, e.g., the integers in the field $\mathbf{Q}(\sqrt{-5})$. To give an example of a different kind, take R to be the ring generated by x_0, x_1, x_2, x_3 over a field k , with the defining relation

$$(2) \quad x_0 x_1 = x_2 x_3.$$

Here $x_1 \mid x_2 x_3$, but $x_1 \nmid x_2$, $x_1 \nmid x_3$, so x_1 is not prime, but it is clearly an atom.

Regarding atomicity, this is a finiteness condition which clearly holds in every Noetherian domain. More generally, it holds in every integral domain with ascending chain condition on principal ideals, ACC_1 for short. Conversely, every UFD satisfies ACC_1 , but there are atomic domains not possessing ACC_1 (notwithstanding an assertion to the contrary in Proposition 1.1 of [12]), e.g., the ring of all polynomials in x and y with rational coefficients, but where $x^r y^s$ has an integral coefficient whenever $rs = 0$. On the other hand, the Noetherian condition is not necessary in a UFD, as we have seen, but as a rule this is the important case for algebraic geometry.

A less obvious factoriality criterion is obtained by using prime ideals. An ideal \mathfrak{p} in a ring R is said to be *Prime* if R/\mathfrak{p} is an integral domain. E.g., in an integral domain a principal ideal (p) is prime precisely when p is 0 or a prime element. Another way of describing a prime ideal is as an ideal whose complement is **multiplicative**, i.e., a nonempty multiplicatively closed set.

We recall that given any subset S of a ring R , and any ideal \mathfrak{a} disjoint from S , the standard application of Zorn's lemma produces an ideal \mathfrak{p} containing \mathfrak{a} and maximal subject to the condition $\mathfrak{p} \cap S = \emptyset$. Moreover, if S is multiplicative, \mathfrak{p} is prime, as is easily checked (and well known). Then we have the following characterization of UFD's in terms of prime ideals [28]:

THEOREM 2. *An integral domain is a UFD if and only if every nonzero prime ideal contains a prime element.*

We recall the essence of the proof. If R is a UFD and $\mathfrak{p} \neq 0$ a prime ideal, let \mathfrak{p} contain $a = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$, where the p_i are primes, then $p_i \in \mathfrak{p}$ for some $i = 1, \dots, r$. Conversely, assume the condition and let S be the set of all products of primes. Then S is multiplicative and moreover, it is **saturated**, i.e., any factor of an element of S is itself in S . If there is a non-unit c not in S , then $(c) \cap S = \emptyset$, so a maximal ideal \mathfrak{p} containing c and disjoint from S exists; it must be prime and by hypothesis contains a prime element, which contradicts the fact that all these elements lie

in S . Now the usual proof of Theorem 1 shows that a product of primes is necessarily unique.

In particular we have the

COROLLARY 1. *In a UFD, every minimal non-zero prime ideal is principal.*

In a Noetherian domain the converse holds, i.e., a Noetherian domain in which every minimal non-zero prime ideal is principal is a UFD; this follows from Theorem 2, because in this case every non-zero prime ideal contains a minimal non-zero prime ideal. But this is a non-trivial result, the consequence of Krull's 'principal ideal theorem' (cf. [28], where the latter is described as "probably the most important single theorem in the theory of Noetherian rings").

To give a geometrical illustration, consider a twisted cubic. This cannot be obtained as the complete intersection of two surfaces in 3-space, for if the surfaces had degrees m and n , then $mn = 3$ and so m or n is 1 and the cubic would be plane. In fact, as is well known, a twisted cubic can be obtained as the intersection of three suitable quadrics, or as the intersection of two quadrics with a common generator, if we ignore the generator.

Geometrically, any non-degenerate quadric in complex projective 3-space can be brought to the form $x_0x_1 = x_2x_3$. The ring of functions on this quadric is the ring $A = \mathbb{C}[X_0, X_1, X_2, X_3]/(f)$, where $f = X_0X_1 - X_2X_3$. This ring is an integral domain, but not a UFD, as we saw earlier. Now a minimal non-zero prime ideal of A corresponds to a maximal proper subvariety of the quadric, i.e., a curve, and by Th. 2, we cannot always expect this to be given by a single equation; the twisted cubic is a case in point. (This is a slight oversimplification, because subvarieties are actually defined by homogeneous ideals in this case.)

Generally, if k is an algebraically closed field, an algebraic set over k is given by the zeros in k^n of a set of polynomials in X_1, \dots, X_n and the coordinate ring of this algebraic set is

$$A = k[X_1, \dots, X_n]/\mathfrak{a},$$

where \mathfrak{a} is the ideal generated by the given set of polynomials. We have a variety (= irreducible algebraic set) if and only if \mathfrak{a} can be taken to be a prime ideal, and then A is an integral domain. If we take for granted the fact that every maximal subvariety of an n -dimensional variety has codimension 1 (i.e., is $(n-1)$ -dimensional, cf. [32], p. 36), Th. 2, Cor. 1 gives the necessity of the next result; the sufficiency follows by the remark following Th. 2, Cor. 1, because the coordinate ring of a variety is Noetherian.

COROLLARY 2. *The coordinate ring of a variety is a UFD if and only if every subvariety of codimension 1 determines a principal ideal (thus the subvariety is a complete intersection).*

If V is any variety, the set \mathfrak{o}_x of functions defined at a given point x of V is a

local ring, i.e., a ring whose non-units form an ideal (the ideal of functions vanishing at x). If x is a simple point of V , \mathfrak{o}_x is what is called a **regular** local ring, and this is necessarily a UFD. There are several proofs of this fact; for a thorough analysis of the algebraic background, see [28], and for a history of the problem, see [35].

4. Nagata's Theorem. How does one prove that a given ring is a UFD? In the case of \mathbf{Z} we needed the Euclidean algorithm ([19], Book VII, Prop. 1–2) to prove that every atom is prime. For polynomial rings in one variable over a field one can use the same method (introduced by Stevin [40] to find the greatest common divisor of two polynomials), but for more than one variable this method is no longer available (in the next section we shall see why). However, it is still true that a polynomial ring in any number of variables over a field is a UFD. More generally, if R is a UFD, then so is $R[X]$; the proof depends on forming rings of fractions.

Let R be an integral domain with field of fractions K . Given a multiplicative subset S of R^* , write

$$R_S = \{a/s \mid a \in R, s \in S\}.$$

This is again a ring, in fact a subring of K , e.g., $R_{R^*} = K$. Any prime p in R either becomes a unit in R_S or it remains prime, depending on whether or not p divides an element of S . Moreover, any atom in R_S comes from an atom in R . Thus by Th. 1 we obtain:

THEOREM 3. *If R is a UFD and S any multiplicative subset of R^* , then R_S is also a UFD.*

Conversely, we have the following result, first proved by Nagata [34] (for Noetherian domains):

THEOREM 4. *Let R be an atomic integral domain and S a multiplicative subset of R consisting of products of primes. Then if R_S is a UFD, so is R .*

The proof consists roughly in this: every atom of R either divides an atom of S and is then shown to be prime (using the fact that S consists of prime products), or if it divides no element of S it stays an atom in R_S and is then prime because R_S is a UFD (cf. [17], p. 116).

With the help of Th. 4 it is very easy to show that factoriality is preserved by adjunction of indeterminates.

COROLLARY 1. *If R is a UFD, then so is the polynomial ring $R[X]$.*

For if $K = R_{R^*}$ is the field of fractions, then by the Euclidean algorithm, $K[X]$ is a UFD. Now $K[X] = R[X]_{R^*}$ and R^* consists of prime products, because R is a UFD, while Gauss's lemma ensures that any prime of R is still a prime in $R[X]$. Further, $R[X]$ satisfies ACC_1 : one gets a bound on the number of non-unit factors by considering leading terms. Hence by Th. 4, $R[X]$ is a UFD.

Cor. 1 shows (by induction) that the polynomial ring in any finite number of indeterminates over a field is a UFD. To extend the result to infinitely many indeterminates one can proceed as follows.

Let A, B be any rings, such that A is a subring of B . The inclusion $A \subseteq B$ is said to be *inert* if for any $c \in A$ such that $c = ab$, where $a, b \in B$, there exists a unit $u \in B$ such that $au, u^{-1}b \in A$. E.g., any integral domain R is inert in the polynomial ring $R[X]$. Now let R be a ring which is a union of a directed system of subrings R_λ (i.e. any two subrings of the system are contained in a third). If all the R_λ are UFD's and all the inclusions $R_\lambda \subseteq R_\mu$ are inert, then R is a UFD. For any $c \in R$ lies in some R_λ and so has a unique factorization into atoms in R_λ ; moreover any factorization of c in a bigger ring R_μ can by inertia be pulled down to R_λ and so must agree with the factorization already found.

We note that the inertia condition cannot be omitted: the semigroup algebra (over a field) of the additive semigroup of positive rational numbers is not a UFD, but it can be written as the union of a directed system of UFD's.

Now let $\mathcal{X} = (X_i)_{i \in I}$ be an infinite family of indeterminates and (R_λ) the family of rings obtained by adjoining finitely many of the X 's to a given UFD R . The R_λ form a directed system of subrings of $R[\mathcal{X}]$ with inert inclusions, and each is a UFD (by (Cor. 1), hence $R[\mathcal{X}]$ is again a UFD.

As a second application of Th. 4, due to Nagata [34, 37] we show that the coordinate ring of a quadric in more than three dimensions is a UFD.

COROLLARY 2. *Let k be an algebraically closed field of characteristic not two, and $Q(X_1, \dots, X_n)$ a non-degenerate quadratic form in $n \geq 5$ variables. Then $[A = kX_1, \dots, X_n]/(Q)$ is a UFD.*

Proof. We can always write $Q = X_1X_2 + Q_1(X_3, \dots, X_n)$; here Q_1 is irreducible because $n \geq 5$. Writing x_i for the residue class of $X_i \pmod{Q}$, we have $A = k[x_1, \dots, x_n]$. Let S be the multiplicative set generated by X_2 in $k[X_1, \dots, X_n]$ and S' the multiplicative set generated by x_2 in A , then

$$A_{S'} = k[x_2, \dots, x_n][x_2^{-1}] = k[x_2, \dots, x_n]_{S'} = k[X_2, \dots, X_n]_S.$$

Since $k[X_2, \dots, X_n]$ is a UFD, the ring on the right is a UFD by Th. 3. Now x_2 is prime in A because Q_1 is irreducible, so A is a UFD by Th. 4.

Cor. 2 has an interesting geometrical consequence due to Klein. The lines in projective 3-space are described by Plücker coordinates π_{ij} ($i, j = 0, \dots, 3$), subject to the relation (cf. [41], p. 22):

$$(3) \quad \pi_{01}\pi_{23} + \pi_{02}\pi_{31} + \pi_{03}\pi_{12} = 0.$$

Each set of ratios (π_{ij}) satisfying (3) defines a line, so that the set of lines in 3-space may be viewed as a quadric (clearly non-degenerate) in projective 5-space, the **Klein quadric**. Thus the lines in 3-space form an algebraic variety; an algebraic subset of codimension 1 on this variety is called a **line complex**. Now by Th. 4,

Cor. 2, the coordinate ring of the Klein quadric is a UFD and so (by Th. 2, Cor. 2) its subvarieties of codimension 1 are complete intersections. Thus we get:

KLEIN'S THEOREM. *Every irreducible line complex in projective 3-space is given by a single equation in Plücker coordinates (besides (3)).*

5. The lattice method. We now turn to the second method of studying UFD's, the **lattice method**. This starts from quite a different definition of UFD, though of course equivalent to the one given earlier. It leads to other generalizations, and in particular, it does not require the ring to be commutative.

Thus let R be an integral domain (not necessarily commutative). If $c \in R^*$ and

$$(4) \quad c = a_1 \cdots a_r,$$

we consider the sequence of right ideals from R to cR :

$$R \supseteq a_1 R \supseteq a_1 a_2 R \supseteq \cdots \supseteq a_1 \cdots a_r R = cR;$$

with it we associate the corresponding quotients

$$(5) \quad R/a_1 R, a_1 R/a_1 a_2 R \cong R/a_2 R, \cdots, R/a_r R.$$

If we have a second factorization of c :

$$(6) \quad c = b_1 \cdots b_s$$

with quotients $R/b_1 R, \cdots, R/b_s R$, we say that (4) and (6) are **isomorphic** if $r = s$ and there is a permutation $i \mapsto i'$ of $\{1, \cdots, r\}$ such that $R/a_i R \cong R/b_{i'} R$. Now we define a (general) UFD as an atomic integral domain in which any two complete factorizations of a given element are isomorphic.

This provides a definition of UFD for the non-commutative case. Although stated in terms of right modules, it turns out to be left-right symmetric. Further, it reduces to the previous definition in the commutative case; to see this we note that if in a commutative integral domain, $R/aR \cong R/bR$, then aR is the annihilator of R/aR and so $aR = bR$, from which it follows that a and b are associated.

Let R be any ring; a right R -module M is said to be **strictly cyclic** if it can be written as R/cR , where c is a non-zero divisor. We denote by \mathcal{C}_R the category whose objects are all the strictly cyclic right R -modules, while the morphisms are all the homomorphisms between them. The category ${}_R\mathcal{C}$ of strictly cyclic left R -modules is defined correspondingly. Any homomorphism $f: R/aR \rightarrow R/bR$ in \mathcal{C}_R is given by an equation

$$ca = bc',$$

and based on this fact one shows that there is a duality (i.e., a category anti-isomorphism) between \mathcal{C}_R and ${}_R\mathcal{C}$, for any ring R , [13, 17]. We shall call it the **factorial duality** in R ; in particular this shows that

$$(7) \quad R/aR \cong R/bR \text{ if and only if } R/Ra \cong R/Rb,$$

from which the left-right symmetry of the above notion of UFD follows immediately. Let us call two non-zero-divisors a, b of a ring R **similar** if $R/aR \cong R/bR$ [20, 24]. By (7) this notion is left-right symmetric; the corresponding notions for zero-divisors are distinct, as examples by Fitting [20] show.

Earlier we saw that in a commutative ring two elements are similar precisely when they are associated; in general there is no such simple criterion. Some equivalent conditions are given in

THEOREM 5. *Let a, a' be two non-zero-divisors in a ring R . Then the following three conditions are equivalent:*

- (i) a and a' are similar,
- (ii) the matrices $\begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 0 & a' \end{pmatrix}$ are associated in R_2 ,
- (iii) there exist mutually inverse 2×2 matrices μ with a in the $(1, 1)$ -position and μ^{-1} with a' in the $(2, 2)$ -position.

For a proof see [17], p. 124 f. ((i) \Leftrightarrow (ii) was proved in [20] and (i) \Leftrightarrow (iii) in [9]). Here are some examples of non-commutative UFD's.

1. The ring of integral quaternions (a rational quaternion is said to be **integral** if its coefficients are integers or halves of odd integers).

2. The ring of linear differential operators. Let $k = C(t)$ be the field of rational functions in a single variable and D an indeterminate over k with the commutation rule

$$(8) \quad Df = fD + f', \text{ where } f' = df/dt.$$

The skew polynomials $\sum f_i D^i$ ($f_i \in k$) with multiplication according to (8) form a UFD. This, probably one of the first examples, was established by Landau [31] and Loewy [33]. Two polynomials in D which are similar in the sense explained above define differential equations which are equivalent in the sense of Poincaré.

3. Free associative algebras [13, 17]. Every free associative algebra $k\langle x_1, \dots, x_n \rangle$ over a field is a UFD; as an example of a non-trivial factorization in the free algebra $k\langle x, y \rangle$ we have

$$xyx + x = x(yx + 1) = (xy + 1)x,$$

and as is easily seen, $xy + 1$ and $yx + 1$ are similar atoms.

4. Group algebras of free groups [13, 17].

5. Free products of skew fields [13].

The definition given at the beginning of this section suggests a way of proving a ring to be a UFD:

THEOREM 6. *An integral domain R is a UFD whenever for each $c \in R^*$ the set $L(cR, R)$ of principal right ideals between R and cR forms a modular lattice of finite length of the lattice of all right ideals of R .*

For then we can apply the Jordan-Hölder theorem for modular lattices [4, 10]. An example of a modular lattice is the lattice of all right ideals of a ring R ; this is the lattice of all submodules of R regarded as a right R -module. The lattice of all submodules of any module is modular (hence the name). Therefore in a principal ideal domain (i.e., an integral domain in which every left or right ideal is principal) the principal right ideals between R and cR form a modular lattice; the *ACC* holds because every right ideal is finitely generated, while the *DCC* follows by the factorial duality, using the fact that R has *ACC* for left ideals. So we obtain

COROLLARY 1. *Every principal ideal domain is a UFD.*

Both the integral quaternions and the ring of linear differential operators are principal ideal domains and therefore are UFD's. The free associative algebra (on more than one free generator) is clearly not a principal ideal domain; so we look for weaker hypotheses from which to deduce Th. 6. Whenever the principal right ideals form a sublattice of the lattice of all right ideals, they form a modular lattice; this is so provided that for any $a, b \in R$ there exist $d, m \in R$ such that

$$aR + bR = dR, \quad aR \cap bR = mR.$$

The first equation leads to the Bezout identity $au + bv = d$ for the greatest common divisor d of a and b , and these rings are called **right Bezout domains**. They are just the integral domains in which every finitely generated right ideal is principal, thus they are not much more general than principal right ideal domains. To get a wider class, let us look at free algebras. Any free associative algebra over a field has the following property [13]:

Every right ideal is free as right R -module, of unique rank.

A ring with this property is called a **free right ideal ring** or **right fir** for short. Left firs are defined similarly and a left and right fir is called a **fir**. E.g., free algebras (over a field), group algebras of free groups, and free products of skew fields are all firs. This is proved by the **weak algorithm**, a generalization of the Euclidean algorithm (to which it reduces in the commutative case). From this point of view the polynomial ring in one variable $k[X]$ is just the free associative algebra on one generator. This explains why the Euclidean algorithm for polynomials in one variable does not extend to more variables: it only applies to free algebras.

Any fir satisfies left and right *ACC*₁ [13, 17]. Moreover, the mapping $(x, y) \mapsto ax - by$ from R^2 to $aR + bR$ defines an exact sequence

$$(9) \quad 0 \rightarrow aR \cap bR \rightarrow R^2 \rightarrow aR + bR \rightarrow 0,$$

which necessarily splits (because $aR + bR$ is free); hence $aR + bR$ and $aR \cap bR$ are principal whenever $aR \cap bR \neq 0$. Thus all the conditions of Th. 6 hold and we find [13, 17]:

COROLLARY 2. *Any (left and right) fir is a UFD.*

Let us define, for any $n \geq 1$, an n -fir as a ring in which every right ideal on at most n generators is free, of unique rank. The notion so defined is left-right symmetric and for larger n we get smaller classes, until we get to **semifirs**, the rings that are n -firs for all n . The 1-firs are just the integral domains, and a 2-fir is a ring in which each 2-generator right ideal is free, of unique rank. Looking more closely to see what was needed to prove Cor. 2, we obtain [9, 17]:

COROLLARY 3. *Every atomic 2-fir is a UFD.*

This generalizes Cor. 2, because every fir is clearly an atomic 2-fir. In the commutative case, every atomic 2-fir is a principal ideal domain, so Cor. 2 and 3 tell us nothing new for commutative rings. But we have seen examples of non-principal firs (free algebras) and there are also atomic 2-firs that are not firs [3, 11]. For example, to obtain a ring R such R^* can be embedded in a group but R cannot be embedded in a (skew) field (Malcev's problem), Bowtell [7] constructs an atomic 2-fir; this ring cannot be a fir because every fir is embeddable in a field [16]. We remark in passing that if R is any atomic 2-fir, then R^* is embeddable in a group [16]; it is not known whether this property is shared by all 2-firs, or by all UFD's.

The problem of unique factorization has also been studied in rings with a set of defining relations of the form $ab = cd$, where a, b, c, d are atoms in the free algebra, by Bokut' [5].

6. The Schreier refinement property. A look at the lattice method of defining UFD's immediately suggests the generalization obtained by giving up atomicity. Let us define an S -ring (for Schreier) as an integral domain in which any two factorizations of any non-zero element have isomorphic refinements.

For the moment let us return to the commutative case; in the presence of ACC_1 , S -rings reduce of course to UFD's, but in general these classes are distinct. In fact there is an intermediate class, the HCF -rings: an HCF -ring is an integral domain in which any two elements have a highest common factor. In terms of the divisibility group D of a ring (section 2) we can say that R is an HCF -ring if and only if D is lattice-ordered, while R is an S -ring precisely if D has the (m, n) -interpolation property, for all m, n : given $x_1, \dots, x_m, y_1, \dots, y_n \in D$, if $x_i \leq y_j$ ($i = 1, \dots, m$, $j = 1, \dots, n$), then there exists $z \in D$ such that $x_i \leq z \leq y_j$ (all i, j). This is actually a consequence of the $(2, 2)$ -interpolation property [4, 12].

Clearly we have the implications

$$\text{UFD} \Rightarrow \text{HCF-ring} \Rightarrow \text{S-ring},$$

and neither of these arrows can be reversed [12]. Moreover, any HCF -ring is integrally closed (in its field of fractions), but this need not be true of S -rings, as is shown by the following example, due (independently) to G. M. Bergman and M. Kneser (unpublished):

Let F be a field with a proper algebraic extension E , and consider the ring of

all formal power series $a_0 + \sum_1^\infty a_i x^{\lambda_i}$, where $a_0 \in F$, $a_i \in E$ ($i > 0$) and (λ_i) is a sequence of positive rational numbers tending to infinity. Then R is an S -ring, but not integrally closed.

By an argument somewhat analogous to the proof of Th. 4 one shows that if R is an integrally closed S -ring, then so is $R[X]$ (cf. [12]). Here the hypothesis of integral closure cannot be omitted; in fact if $R[X]$ is an S -ring, it is easy to see that R must be an integrally closed S -ring. Thus it is more natural to confine attention to integrally closed S -rings. These rings are studied in [12], where they are called **Schreier rings**.

Turning to the non-commutative case, we note that every 2-fir is an S -ring. It is not difficult to give examples of non-atomic 2-firs: apart from the commutative examples there is the group algebra of a free product of copies of the additive group of rational numbers; this is a non-Ore semifir which is non-atomic. It seems more difficult to produce examples of non-Ore semifirs (or even 2-firs) that are atomless.

The commutative case suggests that there should be a condition analogous to integral closure which plays a part in the study of general S -rings, but it is not clear what form this condition should take, or indeed, what its precise role would be.

For other studies of infinite factorizations see [1, 3, 25, 27].

7. Special cases in the non-commutative theory. In a commutative UFD the principal ideals form a lattice; more generally this is so (by definition) in a commutative integral domain with highest common factors (HCF) and least common multiples (LCM). These are the HCF -rings we met in section 6; in fact it is enough to assume the existence of a HCF for each pair of elements, or equivalently assume the existence of a LCM for each pair [12]. Curiously enough, this symmetry disappears when we consider individual pairs: in an integral domain, any pair of elements having an LCM also has an HCF , but the converse need not hold (consider the elements 2 and $2x$ in the subring of the polynomial ring $\mathbf{Z}[X]$ consisting of all polynomials with even coefficient of X).

If R is an HCF -ring and K its field of fractions, then the principal fractional ideals in K form a group under multiplication, and this group structure is compatible with the ordering by inclusion. Thus we have a lattice-ordered group; it is well known that such a group is distributive as a lattice ([4], p. 292). In particular, in a Bezout domain R , for any $c \in R^*$ the set $L(cR, R)$ of principal ideals between R and cR is a distributive lattice. For non-commutative rings this need not be so, even in the case of 2-firs, where $L(cR, R)$ is a sublattice of the lattice of all right ideals. Let us say that an integral domain R has a **distributive factor lattice** if for each $c \in R^*$ the set $L(cR, R)$ is a distributive sublattice of the lattice of all right ideals of R . By the factorial duality the notion so defined is left-right symmetric; moreover, any ring with a distributive factor lattice is a 2-fir.

A principal ideal domain has a distributive factor lattice if and only if every (left or right) ideal is two-sided [17] and this is a fairly stringent requirement. It

is therefore of interest that among general 2-firs quite a wide class of rings have a distributive factor lattice, e.g., free associative algebras and group algebras of free groups. This follows from some technical results of G. M. Bergman proved in [3, 17]. These results show more generally that a 2-fir defined over a field k , which remains a 2-fir under all field extensions, has a distributive factor lattice.

In an atomic 2-fir with distributive factor lattice, the factors of a given element form a distributive lattice of finite length. These lattices have been described in terms of partially ordered sets [3, 4, 17]; to be precise, the categories of finite distributive lattices and homomorphisms, and finite partially ordered sets and order-preserving mappings are dual to each other via the functor $\text{Hom}(-, 2)$, where 2 is the 2-element lattice resp. ordered set. This description has been used by Bergman to study the possible factorizations that can occur. For example, in a commutative UFD, the only distributive lattices which can be realized in this way are direct products of chains; the corresponding partially ordered sets are disjoint unions of chains. However, in a free associative algebra, every finite distributive lattice can be realized as a lattice of factors in this way. The simplest case not occurring in commutative rings is the partially ordered set $\begin{smallmatrix} \circ & & \circ \\ & \searrow & \swarrow \\ & \circ & \end{smallmatrix}$, with the corresponding lattice $\begin{smallmatrix} \square \\ \diagup \end{smallmatrix}$. It is the factor lattice of the element $x(x+1)y$ in the free algebra on x and y (cf. [3, 17]).

We can specialize 2-firs still further by requiring the set $L(cR, R)$ of principal right ideals to be a chain. Any element c with this property is said to be **rigid**, and an integral domain in which all non-zero elements are rigid is called a **rigid domain**. A commutative domain is rigid precisely if it is a valuation ring (by definition of the latter), and a rigid commutative UFD is a discrete valuation ring.

Among non-commutative rings a typical example is the ring of formal power series in several non-commuting indeterminates over a field: $k \llbracket x_1, \dots, x_n \rrbracket$ [17]. More generally, an integral domain is rigid if and only if it is a 2-fir and a local ring. An element c in an atomic 2-fir is rigid whenever all the factors of c in a complete factorization generate a proper ideal [29, 17].

A right discrete valuation ring may be defined as an integral domain R with an atom p such that every non-zero right ideal has the form $p^n R$ and $\bigcap p^n R = 0$. Then a rigid UFD is a right discrete valuation ring if and only if it contains a non-unit c such that cR meets every non-zero right ideal of R non-trivially [17].

8. Remarks on the definition of non-commutative UFD. In some respects the definition of non-commutative UFD given in section 5, is not entirely satisfactory: there is no analogue to Nagata's theorem (Th. 4). Any reasonable analogue should enable one to prove that the free algebra $\mathbb{Z}\langle x_1, \dots, x_n \rangle$ over the integers is a UFD, but this is not the case according the above definition, as the factorizations

$$xyx + 2x = x(yx + 2) = (xy + 2)x$$

show. They are complete factorizations of $xyx + 2x$, but $xy + 2$ is not similar to $yx + 2$.

In order to describe the various possibilities that can arise, let us assume that we have an equivalence relation q defined on the set of non-zero-divisors in each ring such that

- E. 1. If $aq a'$ and a is an atom, then so is a' ,
 E. 2. In a commutative integral domain, $aq a'$ if and only if a is associated to a' .

We shall call R a q -UFD if every element not zero or a unit has a complete factorization into atoms, and given any two such factorizations of the same element:

$$c = a_1 \cdots a_r = b_1 \cdots b_s,$$

we have $r = s$ and there exists a permutation $i \mapsto i'$ of $\{1, \dots, r\}$ such that $a_i q b_{i'}$. For example, the class of UFD's defined in section 5 may now be described more accurately as **similarity-UFD's**. As we remarked earlier, $\mathbb{Z}\langle x, y \rangle$ is not a similarity-UFD, and we therefore try to find a wider equivalence q than similarity, for which this ring is a q -UFD. A number of different choices for q have been proposed; all are wider than similarity and in addition to E. 1-2 satisfy

- E. 3. In an atomic 2-fir, q reduces to similarity.

Their usefulness depends on the ease with which q can be checked; apart from this, the main requirement is that the property of being a q -UFD should be reflected by taking rings of fractions (i.e., Nagata's theorem). To describe this property we must first define primes in non-commutative rings.

An element c of a ring R is said to be **invariant** if c is a non-zero-divisor such that $cR = Rc$. E.g., in a commutative integral domain every non-zero element is invariant. In general the condition on c just states that the left multiples of c are the same as the right multiples; for an invariant element c we can therefore write $c \mid b$ without ambiguity to indicate that b is a multiple of c . Now a **prime** is defined as an invariant non-unit p in R , such that

$$p \mid ab \text{ implies } p \mid a \text{ or } p \mid b.$$

Clearly any product of invariant elements is invariant; thus if S is a multiplicative set consisting of prime products, then every $s \in S$ is invariant and hence, for each $a \in R$, there exists $a' \in R$ satisfying $as = sa'$. This shows that the pair R, S satisfies the Ore right multiple condition, and so S is a right denominator set [15], which can be used to form the ring of fractions R_S . We note that since S consists of non-zero-divisors, the natural homomorphism $\lambda: R \rightarrow R_S$ is injective, so we may take R to be embedded as subring in R_S .

Suppose now that we have an equivalence q defined on each ring R , satisfying E. 1-3 and moreover,

E. 4. Let R be a ring and S a multiplicative subset consisting of prime products. If $a, a' \in R^*$ are such that $a \, q \, a'$ in R_S , then $a \, q \, a'$ in R .

With the help of this condition it is possible to establish an analogue of Nagata's theorem. Thus let R be an atomic integral domain and S a multiplicative set consisting of prime products, such that R_S is a q -UFD, where q is an equivalence relation satisfying E. 1–4. Then R is also a q -UFD; for given two atomic factorizations

$$(10) \quad c = a_1 \cdots a_r = b_1 \cdots b_s,$$

if one of $a_1, \dots, a_r, b_1, \dots, b_s$ is a prime, it is a left factor of c and so may be taken to be a_1 say. Thus $a_1 \mid b_1 \cdots b_s$ and by primeness, $a_1 \mid b_i$ for some i ; since b_i is an atom, it must be associated to a_1 and so b_i is also prime. But then we can divide (10) by a_1 and use induction on $r + s$ to complete the proof that every element has a unique factorization into atoms. We may therefore assume that no a_i or b_j is prime; then it follows that the a_i and b_j cannot divide any element of S . Going over to R_S we find that each a_i, b_j is still an atom in R_S . Since R_S is a q -UFD, $r = s$ and there is a permutation $i \mapsto i'$ of $1, \dots, r$ such that $a_i \, q \, b_{i'}$ in R_S , and by E. 4, $a_i \, q \, b_{i'}$ in R . Thus we have proved:

THEOREM 7. *Let q be an equivalence on rings satisfying E. 1–4. If R is an atomic integral domain and S a multiplicative set consisting of prime products such that R_S is a q -UFD, then R is a q -UFD.*

What was said earlier shows that similarity is an equivalence relation satisfying E. 1–3 but not E. 4. Since we are looking for an equivalence q wider than similarity, two elements a, b defining isomorphic modules $R/aR, R/bR$ will lie in the same q -class, so that q can be described in terms of the category \mathcal{C}_R of strictly cyclic modules. Brungs in [8] defines a preordering ' $<$ ' on R by putting $a < b$ whenever there is an injective homomorphism $R/aR \rightarrow R/bR$. The associated equivalence: ' $a < b$ and $b < a$ ' is called (**right**) **subsimilarity**. It satisfies E. 1–3 and in place of E. 4 satisfies an analogous condition (with a rather more complicated definition of prime), which enables one to prove, e.g., that any free algebra over \mathbb{Z} is a subsimilarity-UFD.

A still wider notion of equivalence was introduced in [14]: We again define a preordering by putting a before b whenever there is a monomorphism $R/aR \rightarrow R/bR$ (i.e., a right cancellative map in \mathcal{C}_R); the associated equivalence is called **right monosimilarity**. This satisfies E. 1–3 and also E. 4 if we limit ourselves to multiplicative sets S in the centre of the ring. In this way we reach the notion of a **right monosimilarity-UFD**; dually one can define **right episimilarity-UFD**'s, on replacing mono- by epimorphisms, but by the factorial duality this is the same as a left monosimilarity-UFD [14].

Still wider notions of equivalence are possible. Let us call two elements a, b of a ring R **left coprime** if they have no common left factor apart from units, and define **right coprime** similarly. A relation

$$(11) \quad ab' = ba'$$

is said to be **coprime** if a, b are left coprime and a', b' right coprime. It can be shown ([17] p. 126) that in a 2-fir two elements a, a' are similar if and only if they can be put in a coprime relation (11). This shows that the relation between a and a' , expressed by a coprime equation (11) in a 2-fir, is an equivalence. In general rings this is not so, but we can construct an equivalence as follows. Let us say that a, a' (in that order) are **perspective** if they can be put in a coprime relation (11); two elements a, a' will be called **projective** if there is a chain $a_0 = a, a_1, \dots, a_n = a'$ such that for $i = 1, \dots, n$ either a_{i-1}, a_i or a_i, a_{i-1} are perspective. Then projectivity is an equivalence, in fact it is the equivalence 'generated' by perspectivity, and by what has been said, it reduces to similarity in 2-firs. This relation has been studied by Beauregard [2] for certain classes of rings. Let us define a **weak HCF-ring** as an integral domain R such that for each $c \in R^*$ the set $L(cR, R)$ is a modular lattice, relative to the ordering by inclusion. By the factorial duality this notion is left-right symmetric, and in the commutative case, it reduces to the notion of *HCF-ring* considered in section 6. Beauregard [2] shows that every atomic weak *HCF-ring* is a projectivity-UFD. Now it is not hard to verify that projectivity is an equivalence satisfying E. 1-4, therefore Th. 7 applies in this case.

To give an example of a ring which definitely falls outside all these definitions of UFD, let us take the Weyl algebra, i.e., the ring A generated by elements x, y with the defining relation $xy - yx = 1$ over a field k of characteristic not 2. This ring is a Noetherian domain and hence has a skew field of fractions. Let S be the set of all non-zero polynomials in y , then S is a right denominator set and the ring A_S of fractions is a principal ideal domain and hence a (similarity-)UFD. However, we have the following factorizations in A :

$$xyx + x = (xy + 1)x = x^2y.$$

It is easily checked that $x, y, xy + 1$ are atoms, so not even the number of factors in a complete factorization is constant.

9. Factorizing zerodivisors. So far we have confined ourselves almost entirely to integral domains. A very similar theory is possible for the factorization of non-zero-divisors in general rings, but this is of less interest and so has not received as much attention. For 'full' matrices over firs there is a fairly satisfactory theory (cf. [17], ch. 5, and, for an application, [16]). The corresponding theory for rectangular matrices still faces difficulties, in that not all complete factorizations of a given matrix have the same number of factors [17].

Finally there is the problem of factorizing zero-divisors. A definition of commutative unique factorization ring (with zero-divisors) has been given by Fletcher [21, 22], who shows that the unique factorization rings so defined are just the finite direct products of UFD's and 'special' principal ideal rings (i.e., homomorphic images of discrete valuation rings). The main difficulty in factorizing zero-divisors

is that one cannot expect uniqueness; nevertheless legitimate questions can be asked, as is shown by the theorem on the diagonal reduction of matrices over a principal ideal domain, which may be regarded as the prototype of a unique factorization theorem for this case.

References

1. R. A. Beauregard, Infinite primes and unique factorization in a principal right ideal domain, *Trans. Amer. Math. Soc.*, 141 (1969) 245–254.
2. ———, Right LCM-domains, *Proc. Amer. Math. Soc.*, 30 (1971) 1–7.
3. G. M. Bergman, Commuting elements in free algebras and related topics in ring theory, Thesis, Harvard University, 1967.
4. G. Birkhoff, *Lattice theory*, 3rd ed. (AMS, Providence 1967).
5. L. A. Bokut', Factorization theorems for certain classes of rings without zero-divisors (Russian) I, *Algebra i Logika* 4, No. 4 (1965) 25–52; II, *ibid.* No. 5 (1965) 17–46.
6. N. Bourbaki, *Algèbre commutative*, ch. 7 (Hermann, Paris 1965).
7. A. J. Bowtell, On a question of Malcev, *J. Algebra*, 6 (1967) 126–139.
8. H. H. Brungs, Ringe mit eindeutiger Faktorzerlegung, *J. Reine Angew. Math.*, 236 (1969) 43–66.
9. P. M. Cohn, Noncommutative unique factorization domains, *Trans. Amer. Math. Soc.*, 109 (1963) 313–331; correction 119 (1965) 552.
10. ———, *Universal algebra*, Harper & Row, New York, London, Tokyo, 1965.
11. ———, Some remarks on the invariant basis property, *Topology*, 5 (1966) 215–228.
12. ———, Bezout rings and their subrings, *Proc. Cambridge Phil. Soc.*, 64 (1968) 251–264.
13. ———, Free associative algebras, *Bull. London Math. Soc.*, 1(1969) 1–39.
14. ———, Factorization in general rings and strictly cyclic modules, *J. Reine Angew. Math.*, 239/40 (1970) 185–200.
15. ———, Rings of fractions, this MONTHLY, 78 (1971) 596–615.
16. ———, The embedding of firs in skew fields, *Proc. London Math. Soc.*, (3) 23 (1971) 193–213.
17. ———, *Free Rings and their relations*, Academic Press, London & New York, 1971.
18. R. Dedekind, Über die Theorie der ganzen algebraischen Zahlen, Vieweg, Braunschweig, 1964.
19. Euclid, *Elements* (~300).
20. H. Fitting, Über den Zusammenhang zwischen dem Begriff der Gleichartigkeit zweier Ideale und dem Äquivalenzbegriff der Elementarteilertheorie, *Math. Ann.*, 122 (1936) 572–582.
21. C. R. Fletcher, Unique factorization rings, *Proc. Cambridge Phil. Soc.*, 65 (1969) 579–583.
22. ———, The structure of unique factorization rings, *Proc. Cambridge Phil. Soc.*, 67 (1970) 535–540.
23. A. Grothendieck, *Éléments de géométrie algébrique* (PUF, Paris 1960).
24. N. Jacobson, *Theory of rings*, AMS, Providence, 1943.
25. A. V. Jategaonkar, A counter-example in homological algebra and ring theory, *J. Algebra*, 12 (1969) 418–440.
26. R. E. Johnson, Unique factorization in a principal right ideal domain, *Proc. Amer. Math. Soc.*, 16 (1965) 526–528.
27. ———, Unique factorization monoids and domains, *Proc. Amer. Math. Soc.*, 28 (1971) 397–404.
28. I. Kaplansky, *Commutative rings*, Allyn & Bacon, Boston, 1970.
29. E. G. Koševoi, On the multiplicative semigroup of a class of rings without zero-divisors (Russian), *Algebra i Logika* 5, No. 5 (1966) 49–54.
30. W. Krull, *Idealtheorie*, *Ergeb. d. Math.* vol. 4, 3, Springer, Berlin, 1935.

31. E. Landau, Ein Satz über die Zerlegung homogener linearer Differentialausdrücke in irreduzible Faktoren, *J. Reine Angew. Math.*, 124 (1902) 115–120.
32. S. Lang, *Introduction to algebraic geometry*, Interscience, New York, 1958.
33. A. Loewy, Über reduzible homogene Differentialausdrücke, *Math. Ann.*, 56 (1903) 549–584.
34. M. Nagata, A remark on the unique factorization theorem, *J. Math. Soc. Japan*, 9 (1957) 143–145.
35. ———, *Local rings*, Interscience, New York — London, 1962.
36. P. Samuel, Sur les anneaux factoriels, *Bull. Soc. Math. France*, 89 (1961) 155–178.
37. ———, *Anneaux factoriels*, Sao Paulo, 1963.
38. ———, *Lectures on unique factorization domains*, TIFR, Bombay, 1964.
39. ———, Unique factorization, this MONTHLY, 75 (1968) 945–952.
40. S. Stevin, *Arithmétique* (1585, new ed. 1958).
41. B. L. van der Waerden, *Moderne algebraische Geometrie*, Springer, Berlin, 1939.
42. O. Zariski and P. Samuel, *Commutative algebra I*, Van Nostrand, Princeton, 1968.

CONTINUOUS ANALOGUES OF SERIES

R. P. BOAS, JR., Northwestern University and

H. POLLARD, Purdue University

1. Introduction. The present note, which was inspired by [10], arose as an attempt to understand why some infinite series have continuous analogues whereas others do not. The methods of [10] are *ad hoc* and fail to reveal the underlying mechanism.

The notion of continuous analogue is not easy to define precisely; however, given that [1], [6], [7]

$$(1) \quad \sum_{n=-\infty}^{\infty} \frac{\sin^2(c+n)\alpha}{(c+n)^2} = \int_{-\infty}^{\infty} \frac{\sin^2(c+x)\alpha}{(c+x)^2} dx = \pi/\alpha, \quad 0 < \alpha < \pi,$$

where $(\sin^2 \alpha u)/u^2$ is taken as α^2 when $u = 0$, almost anyone would call the integral in (1) a continuous analogue of the series. A less transparent example is (series: see, for example, [5], p. 102; integral: see any book on complex analysis)

$$(2) \quad \sum_{n=1}^{\infty} \frac{\sin(n - \frac{1}{2})\alpha}{n - \frac{1}{2}} = \int_0^{\infty} \frac{\sin \alpha x}{x} dx = \frac{\pi}{2} \operatorname{sgn} \alpha, \quad |\alpha| < 2\pi,$$

where the analogy seems flawed; but it is improved if we rewrite (2) as

$$(3) \quad \sum_{n=-\infty}^{\infty} \frac{\sin(n - \frac{1}{2})\alpha}{n - \frac{1}{2}} = \int_{-\infty}^{\infty} \frac{\sin(x - \frac{1}{2})\alpha}{x - \frac{1}{2}} dx.$$

Recently Pollard and Shisha [10] observed that although the binomial series

$$(4) \quad (1 + e^{it})^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} e^{int}, \quad |t| < \pi, \quad -\pi < t < \pi; \quad \alpha > -1$$

does not have a continuous analogue — that is, we do not get a correct result by replacing n in (4) by x and replacing summation by integration — it does if we first extend the sum in (4) over $(-\infty, \infty)$. This does not change the series because the added terms are all zero, and it turns out that, in fact,

$$(5) \quad \int_{-\infty}^{\infty} \binom{\alpha}{u} e^{iut} du = \sum_{n=-\infty}^{\infty} \binom{\alpha}{n} e^{int} = (1 + e^{it})^\alpha$$

for $\alpha > -1$ and $|t| < \pi, -\pi < t < \pi$.

Thus a series that does not have a continuous analogue may acquire one if we write a different but equivalent formula for the terms of the series.

As another example, it originally puzzled us that the binomial series (4) has a continuous analogue whereas the equally natural

$$(6) \quad (1 - e^{it})^{-\alpha} = \sum_{n=0}^{\infty} \binom{\alpha + n - 1}{n} e^{int}, \quad \alpha < 1, \quad 0 < t < 2\pi$$

does not, even if we extend the sum over $(-\infty, \infty)$.

If, however, we write (6) in the equivalent form

$$(7) \quad (1 - e^{it})^{-\alpha} = \sum_{n=0}^{\infty} \binom{\alpha + n - 1}{n} \frac{\sin \pi(n + \alpha)}{\sin \pi\alpha} e^{in\pi} e^{int}, \quad 0 < t < 2\pi,$$

it is true (as we shall see) that also

$$(8) \quad (1 - e^{it})^{-\alpha} = \int_{-\infty}^{\infty} \binom{\alpha + u - 1}{u} \frac{\sin \pi(u + \alpha)}{\sin \pi\alpha} e^{iu\pi} e^{iut} du.$$

Continuous analogues of series are of interest in physics (cf. [2]), where one often attempts to deal with an intractable sum by replacing it by the corresponding integral. In fact, the sum in (1) does arise in physics and was “approximated” by the integral before it was realized that the approximation is exact. (See [1], [6], [7].)

2. A general formula. We shall need the notation, but only the simplest theorems, of the theory of Fourier transforms. (Everything that we use is in [11].) Our notation is

$$(9) \quad g(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} e^{-iux} G(u) du, \quad G(u) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} g(x) e^{iux} dx,$$

and we shall use other letters in the same way. In all our work, G will be zero outside a certain finite interval, and we shall suppose that $|G|$ is integrable.

Suppose now that $G(x) = 0$ for x outside an interval $(t - 2\pi, t + 2\pi)$, with $G(x) \rightarrow 0$ as $x \rightarrow t \pm 2\pi$ from inside the interval. Then when n is an integer,

$$\begin{aligned}
 (2\pi)^{\frac{1}{2}}g(n) &= \int_{t-2\pi}^t e^{-inu}G(u)du + \int_t^{t+2\pi} e^{-inu}G(u)du \\
 &= \int_t^{t+2\pi} e^{-inu}\{G(u-2\pi) + G(u)\}du.
 \end{aligned}$$

The numbers on the right are 2π times the Fourier coefficients for the interval $(t, t+2\pi)$ of the function $G(u-2\pi) + G(u)$. Let us suppose that $G(u-2\pi) + G(u)$ satisfies, at t , a sufficient condition for the convergence of its Fourier series to its value at t (for example, it is enough to have the function of bounded variation in a neighborhood of t , its value at t being the average of its right-hand and left-hand limits). Then $\sum g(n)e^{int}$ is a Fourier series and converges to the value at t of the function that generates it; that is,

$$\sum_{n=-\infty}^{\infty} g(n)e^{int} = (2\pi)^{\frac{1}{2}}\{G(t-2\pi) + G(t)\} = (2\pi)^{\frac{1}{2}}G(t).$$

If we replace $G(t)$ by its value from (9), we find

$$(10) \quad \sum_{n=-\infty}^{\infty} g(n)e^{int} = \int_{-\infty}^{\infty} g(x)e^{itx}dx.$$

In practice we usually have $G(x) = 0$ outside a shorter interval (r, s) ; then (10) holds for $s-2\pi < t < r+2\pi$. In particular, it holds for $0 < t < 2\pi$ when $(r, s) = (0, 2\pi)$; and for $-\pi < t < \pi$ when $(r, s) = (-\pi, \pi)$.

Formula (10) is actually a special case of the Poisson summation formula ([11], p. 60; [13], vol. I, p. 68; 2; [8], p. 152), but we do not need the general formula.

3. Examples. We can now produce examples of (10) by looking at functions that are known to have the form

$$(11) \quad g(x) = (2\pi)^{-\frac{1}{2}} \int_r^s e^{-iux}G(u)du, \quad s-r < 4\pi$$

(cf. [1], [4], [6], [7]).

Let us first take $G(u) = e^{-icu}$ on $(-\alpha, \alpha)$ and $G(u) = 0$ for $|u| > \alpha$, where $0 < \alpha < \pi$. Then

$$(12) \quad g(x) = (2\pi)^{-\frac{1}{2}} \int_{-\alpha}^{\alpha} e^{-icu-ixu}du = 2(2\pi)^{\frac{1}{2}} \frac{\sin(c+x)\alpha}{c+x},$$

with the convention that $u^{-1} \sin \alpha u = \alpha$ when $u = 0$. Hence we have (10) for $\alpha-2\pi < t < \alpha+2\pi$, that is

$$\sum_{n=-\infty}^{\infty} \frac{\sin(c+n)\alpha}{c+n} e^{int} = \int_{-\infty}^{\infty} \frac{\sin(c+x)\alpha}{c+x} e^{itx}dx,$$

provided that $0 < \alpha < \pi$, $-\pi < t < 2\pi$. In particular, we can take $t = 0$ and then

$$\sum_{n=-\infty}^{\infty} \frac{\sin(c+n)\alpha}{c+n} = \int_{-\infty}^{\infty} \frac{\sin(c+x)\alpha}{c+x} dx = \pi;$$

since the integrand is an odd function the formula can be written in the form

$$\sum_{n=-\infty}^{\infty} \frac{\sin(c+n)\alpha}{c+n} = \int_{-\infty}^{\infty} \frac{\sin(c+x)\alpha}{c+x} dx = \pi \operatorname{sgn} \alpha, \quad |\alpha| < \pi.$$

For $c = \frac{1}{2}$ we have (2) and for $c = 0$ we have

$$\sum_{n=-\infty}^{\infty} \frac{\sin n\alpha}{n} = \int_{-\infty}^{\infty} \frac{\sin x\alpha}{x} dx = \pi \operatorname{sgn} \alpha, \quad |\alpha| < \pi;$$

remembering that the term with $n = 0$ is to be interpreted as α , we have a symmetrical version of a familiar Fourier expansion.

This discussion can be generalized. The product of two Fourier transforms g_1 and g_2 is again a Fourier transform; if G_1 and G_2 vanish outside $(-\alpha, \alpha)$ and $(-\beta, \beta)$, respectively, then $g_1 g_2$ is the transform of a function vanishing outside $(-\alpha - \beta, \alpha + \beta)$ (actually the function for which $g_1 g_2$ is the transform is the convolution of G_1 and G_2 , but we do not need to know this). Hence if

$$g_1(x) = (2\pi)^{-\frac{1}{2}} \int_{-\alpha}^{\alpha} e^{-iux} G_1(u) du \quad \text{and} \quad g_2(x) = (2\pi)^{-\frac{1}{2}} \int_{-\beta}^{\beta} e^{-iux} G_2(u) du,$$

and $\alpha + \beta < 2\pi$, then

$$\sum_{n=-\infty}^{\infty} g_1(n) g_2(n) e^{int} = \int_{-\infty}^{\infty} g_1(x) g_2(x) e^{itx} dx$$

provided that $\alpha + \beta - 2\pi < t < -\alpha - \beta + 2\pi$.

Taking g_1 as in (12) and $t = 0$, we have in particular

$$(13) \quad \sum_{n=-\infty}^{\infty} g_2(n) \frac{\sin(c+n)\alpha}{c+n} = \int_{-\infty}^{\infty} g_2(x) \frac{\sin(c+x)\alpha}{c+x} dx.$$

If we further specialize by taking g_2 equal to g_1 , we get

$$(14) \quad \sum_{n=-\infty}^{\infty} \frac{\sin^2(c+n)\alpha}{(c+n)^2} = \int_{-\infty}^{\infty} \frac{\sin^2(c+x)\alpha}{(c+x)^2} dx = \frac{\pi}{\alpha}, \quad 0 < \alpha < \pi;$$

this is formula (1) of §1.

4. Binomial series. With Pollard and Shisha, we start from the formula

$$(15) \quad \binom{\alpha}{x} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ixt} (1 + e^{it})^{\alpha} dt, \quad \alpha > -1, \quad -\infty < x < \infty.$$

This is of the form (11) with $r = -\pi$, $s = \pi$; the Fourier series of $(1 + e^t)^{\alpha}$ converges for $|t| < \pi$ by almost any convergence test that we might think of applying.

Then (10) takes the form (5).

Pollard and Shisha also give

$$(16) \quad \sum_{n=-\infty}^{\infty} \binom{\alpha}{n+c} e^{i(n+c)t} = \int_{-\infty}^{\infty} \binom{\alpha}{u+c} e^{i(u+c)t} du = (1 + e^{it})^{\alpha},$$

under the same conditions as (5); this is (10) again with

$$G(t) = \begin{cases} e^{-ict} (1 + e^{it})^{\alpha}, & |t| < \pi; \\ 0, & |t| > \pi. \end{cases}$$

The pair (7), (8) follow in the same way, once we realize that (15) can be transformed into

$$e^{-i\pi x} \frac{\sin \pi(\alpha + x)}{\sin \pi \alpha} \binom{\alpha + x - 1}{x} = \frac{1}{2\pi} \int_0^{2\pi} (1 - e^{it})^{-\alpha} e^{-itx} dt$$

by the usual formulas about the gamma function.

5. A general method. The preceding discussion suggests a general method for constructing continuous analogues of series. Consider a function f defined on the integers, and let $\sum_{n=-\infty}^{\infty} f(n)e^{int}$ be the Fourier series of an integrable function F , so that

$$f(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-inu} F(u) du.$$

There are many conditions that are sufficient for this, for example that $\sum |f(n)|$ converges or that $f(n) \rightarrow 0$ and f is even and convex ([13], vol. I, pp. 183, 326). The function ϕ defined for all real x by

$$\phi(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ixu} F(u) du$$

interpolates f at the integers and has the form (11). Consequently we have

$$\sum_{n=-\infty}^{\infty} f(n)e^{int} = \sum_{n=-\infty}^{\infty} \phi(n)e^{int} = \int_{-\infty}^{\infty} \phi(x)e^{ixt} dx, \quad |t| < \pi.$$

That is, we can construct a continuous analogue of any Fourier series that belongs to a function satisfying the conditions imposed on G in §2. Whether we are willing to regard this as a reasonable analogue seems to depend on whether we can write $\phi(x)$ in a sufficiently recognizable form.

Let us see how the method works out in some specific examples.

We first look for a continuous analogue of the logarithmic series, which we can write in the form

$$it = \sum_{n \neq 0} \frac{(-1)^{n-1}}{n} e^{int}, \quad |t| < \pi.$$

Here

$$\begin{aligned} f(n) &= \frac{i}{2\pi} \int_{-\pi}^{\pi} u e^{-inu} du, \quad n \neq 0; f(0) = 0; \\ F(u) &= iu, \\ \phi(x) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ixu} F(u) du = \frac{1}{x} \left(\cos \pi x - \frac{\sin \pi x}{x} \right). \end{aligned}$$

It is clear that $\phi(n)$ does in fact equal $f(n)$, although there is no really natural compelling analogue of $(-1)^{n-1}/n$, and $\phi(x)$ is perhaps not the most obvious interpolating function. Our continuous analogue is

$$\begin{aligned} \sum_{n \neq 0} \frac{(-1)^{n-1}}{n} e^{int} &= \sum_{-\infty}^{\infty} \frac{1}{n} \left(\cos n\pi - \frac{\sin n\pi}{n\pi} \right) \\ &= \int_{-\infty}^{\infty} \frac{1}{x} \left(\cos \pi x - \frac{\sin \pi x}{\pi x} \right) e^{ixt} dx = it, \quad |t| < \pi. \end{aligned}$$

This seems acceptable; indeed, ϕ is the only possible interpolating function of the form (11).

Now let us look for a continuous analogue of the exponential series,

$$\sum_{n=0}^{\infty} \frac{1}{n!} e^{int} = \exp(e^{it}).$$

Here

$$(17) \quad \phi(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ixu} \exp(e^{iu}) du,$$

and we have $\phi(n) = 1/n!$, $n = 0, 1, 2, \dots$; $\phi(n) = 0$, otherwise;

$$\sum_{n=0}^{\infty} \frac{1}{n!} e^{int} = \int_{-\infty}^{\infty} e^{ixt} dx \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ixu} \exp(e^{iu}) du.$$

This, although formally a pair of analogues, seems unsatisfactory, partly at least because we are conditioned to expect a continuous analogue of $1/n!$ to involve $1/\Gamma(x+1)$, whereas $\phi(x) = 1/\Gamma(x+1)$ only when $x = n$.

We note that a function $\phi(x)$ of the form (18) is easily seen to be (a) bounded on the real axis, (b) of exponential type in the plane, i.e., $|\phi(z)| \leq A e^{\pi|z|}$. But $1/\Gamma(x+1)$ does not satisfy either (a) or (b), for example because $1/\Gamma(-n + \frac{1}{2} + 1) = (-1)^{n-1} \Gamma(n - \frac{1}{2}) \pi \rightarrow \infty$ faster than any e^{Bn} (by Stirling's formula).

We should accordingly like to put $\phi(x)$ into a form that involves $1/\Gamma(x+1)$ explicitly. Now a well-known formula (see [3], vol. 1, p. 13) states that

$$\frac{1}{\Gamma(z+1)} = \frac{1}{2\pi i} \int_{-\infty}^{(0+)} t^{-z-1} e^t dt,$$

where the path of integration can be taken to be the loop extending from $-\infty$ to -1 along the real axis, around zero on the unit circumference, and back to $-\infty$. This yields

$$\phi(x) = \frac{1}{\Gamma(1+x)} - \frac{\sin \pi x}{\pi} \int_1^\infty \frac{e^{-u}}{u^{1+x}} du,$$

and consequently with this $\phi(x)$

$$\sum_{n=0}^{\infty} \frac{1}{n!} e^{int} = \sum_{-\infty}^{\infty} \phi(n) e^{int} = \int_{-\infty}^{\infty} \phi(x) e^{ixt} dx.$$

5. Bessel functions. The generating function for the Bessel functions $J_n(s)$ is

$$\exp(\tfrac{1}{2}s(w - 1/w)) = \sum_{n=-\infty}^{\infty} w^n J_n(s),$$

or, with $w = e^{it}$,

$$(18) \quad \exp(\tfrac{1}{2}s(e^{it} - e^{-it})) = \sum_{n=-\infty}^{\infty} e^{int} J_n(s).$$

Let us look for a continuous analogue of (18). Bessel's integral ([12], p. 19) is

$$J_n(s) = \frac{1}{\pi} \int_0^\pi \cos(n\theta - s \sin \theta) d\theta,$$

when n is an integer. Replacing n by x yields a function of the form (11); unfortunately it is not $J_x(s)$ when x is not an integer, but is known as Anger's function $J_x(s)$ ([12], p. 308). What is true is that ([12], p. 176)

$$J_x(s) + h_x(s) = \frac{1}{\pi} \int_0^\pi \cos(x\theta - s \sin \theta) d\theta,$$

where

$$h_x(s) = \frac{\sin \pi x}{\pi} \int_0^\infty e^{-xr - s \sinh r} dr = J_x(s) - J_x(s),$$

and $h_x(s) = 0$ when x is an integer. Hence an analogue of (18) is

$$\sum_{-\infty}^{\infty} e^{int} (J_n(s) + h_n(s)) = \int_{-\infty}^{\infty} e^{ixt} (J_x(s) + h_x(s)) dx,$$

or alternatively

$$\sum_{-\infty}^{\infty} e^{int} J_n(s) = \sum_{-\infty}^{\infty} e^{int} J_n(s) = \int_{-\infty}^{\infty} e^{ixt} J_x(s) dx,$$

which is no less "natural" than the pair (7), (8).

Added in proof: For similar results see [14].

References

1. A. B. Bhatia and K. S. Krishnan, Light-scattering in homogeneous media regarded as reflexion from appropriate thermal elastic waves, *Proc. Roy. Soc. London, Ser. A.* 192 (1948) 181–194.
2. R. P. Boas, Jr. and C. Stutz, Estimating sums with integrals, *Amer. J. Physics*, 39 (1971) 745–753.
3. A. Erdélyi, et al., *Higher transcendental functions*, McGraw-Hill, New York, etc., 1953.
4. D. Jagerman, Bounds for truncation error of the sampling expansion, *SIAM J. Appl. Math.*, 14 (1966) 714–723.
5. L. B. W. Jolley, *Summation of series*, 2d ed., Dover, New York, 1961.
6. K. S. Krishnan, A simple result in quadrature, *Nature*, 162 (1948) 215.
7. ———, On the equivalence of certain infinite series and the corresponding integrals, *J. Indian Math. Soc.*, (N. S.) 12, (1948) 79–88.
8. L. H. Loomis, *An introduction to abstract harmonic analysis*, Van Nostrand, New York, etc., 1953.
9. B. O. Peirce, *A short table of integrals*, 3d ed., Ginn, Boston etc., 1929.
10. H. Pollard and O. Shisha, Variations on the binomial series, this MONTHLY, 79 (1972) 495–499.
11. E. C. Titchmarsh, *Introduction to the theory of Fourier integrals*, Oxford, 1937.
12. G. N. Watson, *A treatise on the theory of Bessel functions*, 2d ed., Cambridge, 1944.
13. A. Zygmund, *Trigonometric series*, 2d ed., Cambridge, 1959.
14. T. J. Osler, An integral analogue of Taylor's series and its use in computing Fourier transforms, *Math. Comp.*, 26 (1972) 449–460.

ENGLAND WAS LOST ON THE PLAYING FIELDS OF ETON: A PARABLE FOR MATHEMATICS

A. B. WILLCOX, Executive Director, MAA

I am sure that most of you have had the experience at one time or another of discovering, in an unexpected place, an old newspaper, its pages yellowed with age. You may have found that a glance at one of the old news articles jolted your mind into a moment or two of serious reflection on how far we have come since those bygone days. I was rummaging through the attic of my imagination recently when I came across a newspaper dated May 1, 1980, its pages pale with the years not yet lived. A glance at an article I found there jolted my mind into something more than a moment of reflection on where we are going. It was, it seemed to me,

Alfred Willcox received his Yale Ph.D. under Charles Rickart. He served as Instructor through Professor at Amherst College and has held Visiting appointments at the Univ. of Chicago, the Univ. of Uppsala, Sweden, and the Univ. of Wisconsin. His main research interest is functional analysis.

He is presently the Executive Director of the MAA and previously served the MAA on a number of committees, as Second Vice-President, and as Executive Director of CUPM. He is the co-author and editor of the Willcox, Buck, Jacob, Bailey *Calculus Series* (Houghton Mifflin, 1971). *Editor.*

1. A parable for mathematics.

LONDON, MAY 1, 1980. It is only a coincidence, but an interesting one, nevertheless, that the eve of the first Congress of Committees for a New Beginning should fall on the birthday of the Duke of Wellington. An event on which is focused the hope of 60 million inhabitants of this island for a new destiny falls on the day of birth of one of the central figures of that destiny which was once England. England: the hub of a global empire, the dream of a great culture, a nation which inspired love, devotion, sacrifice and pride, an idea which died a quiet death from natural causes on March 23, 1979. This reporter was so intrigued by this confluence of events that he paid a brief visit recently to one of England's last Lords, Jefferey Allyn-Smythe, Lord of Devonshire, Member of Parliament on the day England died.

We chatted, in quite an informal way, about Wellington's England, about the ideals and values which once were the heart and spirit of England, about how these ideals gradually became institutionalized in an Establishment which, once it had become the sole custodian of the spirit and destiny of England, lost its ability to change with the world and adapt to the real needs and desires of the people until it finally died quietly and without fanfare in a vapor of irrelevance.

At one point in our conversation I mentioned those words of Wellington which had once stirred men's hearts, "The battles of England are won on the playing fields of Eton." "That sums it all up," exclaimed Allyn-Smythe, "The rise and fall. It's all there. Wellington, you know, had more in mind than just strong bodies and stout hearts. Eton exemplified all that was grand about England. At Eton the cream of England's youth was prepared, and prepared well, for the kind of service and unswerving loyalty and devotion that built a global Empire and developed a culture as solid as Gibraltar. At least, that was the way it appeared. In actual fact, during the first half of this century Gibraltar began to crumble. It was difficult, the realization that England was after all just a smallish island in a large sea, but we felt that we accommodated to this realization with grace and resolve. What we did not notice was that Eton did not accept the changes which were occurring. The cream of England's youth continued to be prepared, and prepared well, for the kind of service and unswerving loyalty and devotion which was to preserve a myth for two decades. All through the 60's and 70's, Eton and the English Establishment which it represented protected and preserved a skeleton of quality, a dream of greatness, and a pretense of destiny while the world simply lost interest. Education at Eton actually increased in quality and increased in vigor, while Eton itself ignored all the signs that it was becoming irrelevant. The Establishment was still present in all its glory and all of its tradition when the world simply walked away.

"We all know the events which signaled the death of England. In the election of 1979, there simply weren't enough interested British voters to return a Parliament able to govern. The small band of bewildered M.P.'s who were left after the election — there were eleven, weren't there — simply went home. For a year now this tiny

island has coasted along on inertia. The bureaucracy has kept the wheels turning, the managers have kept the store, until the recent riots in a few densely populated areas, and a rising undercurrent of fear across the nation — excuse me, across the land — have led to the formation of the local Committees for a New Beginning which come together tomorrow to fashion a new destiny for a people waiting to become a nation again.

“Whatever that destiny is, it is not Wellington’s. Millions of English hearts swelled with pride when he said, ‘The battles of England are won on the playing fields of Eton.’ I rather think that a Wellington of today might give us words steeped more in irony than pride, offering a challenge instead of proud congratulations: ‘England was lost on the playing fields of Eton.’ We might add, if we profess to have hope, ‘England will be rediscovered in the hearts, and rebuilt by the hands, of Englishmen.’”

I must admit that I find this bit of fantasy, this concocted look into the future, a bit more embarrassing every time I read it. I wrote it in one of those flashes of revelation which occur with brilliant displays of light in the middle of a sleepless night and which turn out to be grotesque Alice-in-Wonderland doggerel in the cold light of the dawn.

Nevertheless, having committed myself to this ridiculous title before really thinking about how I would feel standing here before you reading it, I decided that the only thing to do was to see it through with a straight face. I won’t be hurt if you smile a little, even condescendingly.

I am not embarrassed, however, to suggest that this silly story may indeed be a parable for mathematics. Listen, for example, to this true story, equally insignificant by itself, which ought to give pause to anyone who has his eye open today and who professes an interest and stake in the future of mathematics in our culture.

A friend of mine recently told me that he has been teaching an advanced calculus course after a number of years away from this mainstay of the undergraduate curriculum. He said that he dutifully included in the course a section on line integrals, because it seemed the thing to do and because the topic was contained in the book he was using. After he had given a particularly brilliant lecture on the subject one day, a student came to him and asked why line integrals were contained in the course—what did they have to do with anything outside of mathematics, or even outside of calculus. My friend began describing the usual applications to physics which give line integrals a prominent place in the honor role of “relevant mathematics.” The student interrupted impatiently, “I don’t know anything about physics, and, frankly, I don’t care much about it. I am an economics major. Can you describe any applications of line integrals in the social sciences?” My friend drew a complete blank and fell back on the old defensive position, “I’ll think about it over the weekend and report back to you on Monday.”

He thought several times during the weekend and looked in a number of books which he had in his study, but had no success at all. On Monday he had to confess

that to his knowledge there are no significant applications of line integrals in the social sciences. "Then, why should I, an economics major, be forced to spend a week of my time on line integrals," said the student, "a subject which I can't relate to my own experience and which I find boring?" This sent my friend back to his last line of defense, "But it is beautiful mathematics of great intrinsic worth and appeal." This explanation had always sent his students nodding back to their desks, but not this time. "I am not interested in playing chess, no matter how challenging it is intellectually. I am in college to learn how to do something significant for society, and I am interested in what mathematics can do to help me. I don't have time for chess."

"Do you know," said my friend, "I have heard that many times before and I have devised a hundred devastating rejoinders, but somehow, at that moment, I wanted to shrink to a point and disappear." *I am not interested in playing chess.*

At the end of the last century England literally encircled the globe. It gave to the world the riches of a highly developed culture and a great dream, and in return for this the world repaid England with riches of a more tangible kind. But for all its power and grandeur, England did not solve the pressing problems of hunger, poverty, and lack of shelter in the world and eventually the world simply walked away. That is history. It remains to be seen whether the slide into complete irrelevance is to be continued. I don't really believe that it is, but somehow it doesn't seem impossible.

In 1972, mathematics spans the world of science and much of the world that stands on the edge of science. It has given this world unity, coherence, powerful tools for reason, and much elegance. In return for this, the world has rewarded us richly — we would not be so ungrateful as to deny that. But somehow, we are aware that mathematics has not only failed to solve as many of the great problems as the world expected it to solve, but that it has actually begun to lose contact with much of the world outside its own little island. Is mathematics a game of chess? Will the world lose interest in playing chess?

I would like to share with you a random selection of straws in the wind, drawn from journal and newspaper articles, letters, and other lore, which indicate the directions and the source of my concern, and then to state as succinctly as I can the clear challenge which this wind — still just a gentle breeze — wafts into my inquiring nostrils.

2. Straws in the wind.

STRAW # 1. At its 1970 Annual Meeting in San Antonio the MAA presented a panel discussion on the reports of COSRIMS (Committee on Support of Research in the Mathematical Sciences). During this panel discussion Ralph Boas described COSRIMS as follows [1]:

"What is — or was — COSRIMS? It was a 12-man committee of the National Academy of Sciences under the chairmanship of Lipman Bers, and the name stands for Committee on the

Support of Research in the Mathematical Sciences. Before it was through it had 50 or more collaborators doing things for it, and another 50 or so worked on the CBMS Survey, with John Jewett as executive director, that collected the needed data. The principal product of COSRIMS was simply entitled, "The Mathematical Sciences — a Report"; it was completed, after more than a year of work, at the end of 1967 and issued late in 1968."

In introducing the panel members, the Moderator, Arnold Ross, made the following comments on events which have transpired since the publication of the COSRIMS reports [2]:

"In introducing the COSRIMS Reports, Lipman Bers placed the then current major concerns of our mathematical community into a proper perspective with admirable clarity.

The all pervasive nature of mathematics to which he referred is attested to by the variety of concerns exhibited by these reports. Accepting the premise of the inherent esthetic appeal of mathematics and the demonstrable need for mathematics in the sciences and the professions, the COSRIMS Reports projected an influx of young talent into mathematics on an ever-increasing scale. The problems of education and research, so it seemed, were amenable if only one could throw into the fray sufficient material resources and adequate resources of talent trained to the level of a Ph.D.

However, much has happened since the writing of the reports which put to a severe test the comfortable assurance of our mathematical community.

There has been a strong alienation of young talent away from mathematics. Social unrest has reached the campus and our adequacy as mentors of the younger generation has been vociferously questioned. We have not been able to command material resources which we feel we need for the task of mathematical education at hand. Our expectations for the cream of our mathematical manhood, our young Ph.D.'s, have not been fulfilled — both in regard to what they expected and in regard to what has been expected of them.

As a result of the above disturbing confrontation with the new realities, many voices have been raised to urge a critical reappraisal of our academic responsibilities. A grassroot movement emphasizing the need for a more sensitive response to the needs of our students and to the needs of our young mathematical colleagues, and of our colleagues of all ages in the related sciences has sprung up. The area in grass is not very large as yet but the grass is high. Some of our most accomplished colleagues have joined this movement, thus giving the lie to the all too common heresy that there exists an intrinsic conflict between teaching and research."

STRAW # 2. Earlier, in 1966, Edwin Spanier [3] wrote a memorandum to CUPM expressing some of his concerns about the undergraduate curriculum. In this memorandum, Spanier said,

"It is my contention that the universities and colleges are not doing a good job of educating the undergraduate in mathematics. This is not because we do or don't have certain courses available, but more because of the attitude of the instructors and the emphasis in the courses. Most of the undergraduate mathematics majors are probably reasonably qualified to go to graduate school in mathematics but not for anything else. The demand that existed for mathematics majors in industry seems to have disappeared, and I fear this is because industry has learned that mathematics majors aren't as well trained for their needs as people with other majors.

A mathematics major should learn something about what mathematics is, both in terms of its internal structure and in terms of its relations with other areas. An undergraduate would probably get a more rounded presentation of mathematics in the above sense if he studied

engineering or computer science. This is unfortunate, and we should make a serious effort to change this state of affairs.”

“... We all share responsibility for its state. I don’t know how to change the situation, but I am firmly convinced that if we don’t, we will find ourselves playing a progressively smaller role as that of the engineers and computer scientists grows.....”

STRAW # 3. At that same San Antonio meeting, R. D. Anderson presented a talk “Are There Too Many Ph.D.’s?” [4] He began his talk by answering the question as follows:

“The answer is ‘Not Yet.’ However, a valid interpretation of such an answer is that there may well be “too many Ph.D.’s within a few years. In the author’s judgement, based on the evidence cited below, there should be positions for Ph.D.’s in mathematics for the next two years; however, a large percentage of the available academic positions will be in colleges or universities without Ph.D. programs in mathematics, without research libraries, and with teaching loads which are larger than those prevailing at Ph.D. granting universities. By the fall of 1972 there are likely to be more Ph.D.’s looking for positions than there are (adequately salaried) positions with duties commensurate with Ph.D. level training in mathematics....

“Over the last 25 years, at least, the principal thrust of graduate training in the mathematical sciences has been that of training in core mathematics, chiefly pure mathematics. We have trained Ph.D.’s in our own image, to regard research as the principal purpose of mathematicians and the primary (and almost the only) route to real status in the profession. In a sense we have been fabulously successful. American mathematics has been playing an increasingly important and central role in world mathematics. Research — and good research — has been flourishing as never before. We have inculcated our graduate students with the attitude that teaching was more a means to develop new researchers and a means of support of researchers than an end in itself. And, by and large, we have done little to encourage involvement of our graduate students with applications of mathematics. There is a real need for a rapid change of attitude and action on both of these counts. Many of our young Ph.D.’s will be employed primarily as teachers and many others will need to find positions outside academic life where applications of mathematics will be the basis for their continued employment. Nationally we must alter some of our patterns of graduate education. It does not follow that each university or each graduate student should have a radically different program. It does follow that, statistically, many should.”

Toward the end of his talk, Anderson listed several recommendations based on his observations. Among them were:

- (1) In order to provide the future Ph.D. with necessary options for employment, opportunities for substantial training in applications of mathematics (particularly the new applications) should be offered to (but not necessarily required of) graduate students in core mathematics departments.
- (2) Since academic employment of future Ph.D.’s will be more dependent on teaching performance, greater stress should be placed on training in teaching and for teaching.
- (3) Mathematics departments should actively promote the introduction of more and better ‘service courses, i.e., courses for students in other disciplines. Not only will this accelerate the mathematization of society but it will also increase the demand for mathematics faculty.”

STRAW # 4. Some of my straws fall obliquely on the issue. But sometimes light from the side reveals detail most sharply. In January of 1970 Alvin Weinberg [5] said in an article in SCIENCE magazine, entitled “In Defense of Science”:

"It is incredible, but true, that science and its technologies are today on the defensive. The attack, which is most noticeable in the United States, has been launched on four fronts. First, there are the scientific muckrakers, mostly journalists, who picture the scientific enterprise as being corrupted by political maneuvering among competing claimants for the scientific dollar. Second, there are thoughtful legislators and administrators who see a waning in the relevance of science to the public interest, especially as we address ourselves to grave social questions that are hardly illuminated by science. To deny connection between science and public affairs weakens one of the main arguments for public support of basic science: that out of basic science comes technology, which in turn improves our human condition. Third, there are the many technological critics who urge a slowdown, or at any rate a redirection, of technology because of its detrimental side effects. And finally, there are the scientific abolitionists: the very noisy, usually young, critics who consider the whole scientific-technological, if not rationalistic mode of the past 100 years a catastrophe. To them technology is the opiate of the intellectuals; some of the more extreme would demolish human reason as the ultimate tool for achieving human well-being. The consequence, or perhaps, a further symptom, of all this harassment is a reduction in society's support for science. The U. S. budget for science has fallen from 2.5 percent of the gross national product in 1965 to 2 percent in 1969."

STRAW # 5. This straw has been broken into several pieces. It concerns the current employment situation in mathematics and one interpretation of the meaning of this situation.

An article in the April 26, 1970 Washington Post, [6] entitled "Ph.D. Glut Creating a Jobless U.S. Elite" began:

"I just can't find anything at all," complains a bitter young scientist who won his coveted doctor of philosophy degree at the University of Maryland this past winter and has yet to land a satisfactory job.

"His plight reflects a dramatic development in higher education. The Ph.D. has suddenly ceased being a certain passkey to professional and sometimes financial rewards. The new Ph. D. recipient this year cannot be sure of finding any job that approaches what he had looked forward to during those long, arduous years of postgraduate study."

In an article "Academic Employment Prospects for September 1972," appearing in the February 1972 issue of the NOTICES of the AMS, R. D. Anderson reports the following balance sheet for the academic employment situation in September of 1972. This balance sheet represents an estimate based on recent surveys conducted by the AMS Committee on Employment and Educational Policy.

A BALANCE SHEET FOR ACADEMIC JOBS IN MATHEMATICS

<i>Academic Job Seekers</i>	<i>Academic Jobs Available</i>
1. 900 (new Ph.D.'s not already having jobs)	500 (from survey)
2. 200 (currently professionally unemployed)	100 (death and retirement)
3. 500 (nonretainees)	500 (jobs of nonretainees)
1600 Total jobs seekers	1100 Total jobs
4. -300 Nonpure mathematicians	-300
1300 Pure mathematicians seeking jobs	800 Jobs for pure mathematicians
5. Prospective professionally unemployed pure mathematicians:	
	500 \pm 200.

These two pieces of the straw describe the current uncomfortable employment situation for Ph.D.'s in mathematics. The final piece "explains" this situation in a certain admittedly simplistic and even frivolous way. But the "explanation" is not totally devoid of validity and has a moral! The COSRIMS report, to which I have referred earlier, contains several tables listing the number of earned B.A. and Ph.D. degrees in the mathematical sciences each year from 1954 through 1966. The tables also contain projections of the predicted output of BA's and Ph.D.'s through 1976. The predictions were toned-down versions of projections made by the U.S.O.E. after careful study and the application of the most sophisticated statistical techniques. Now we can compare these predictions with the *actual* output from 1966 to date. These comparisons are contained in the following two charts.

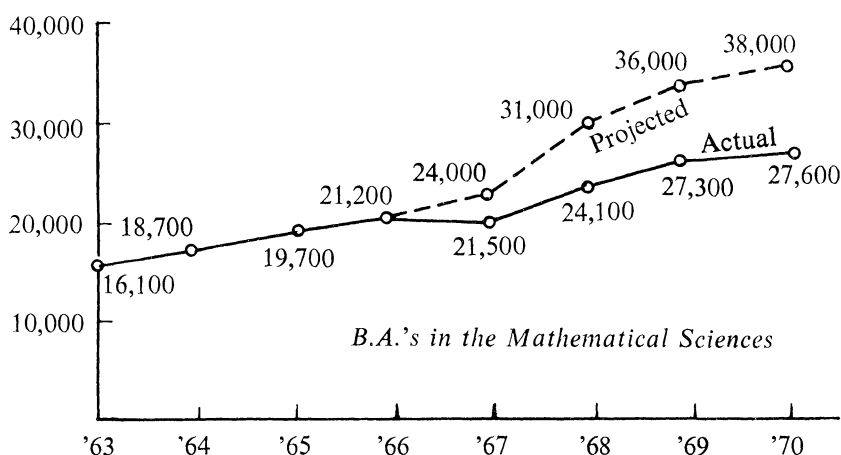


Fig. 1

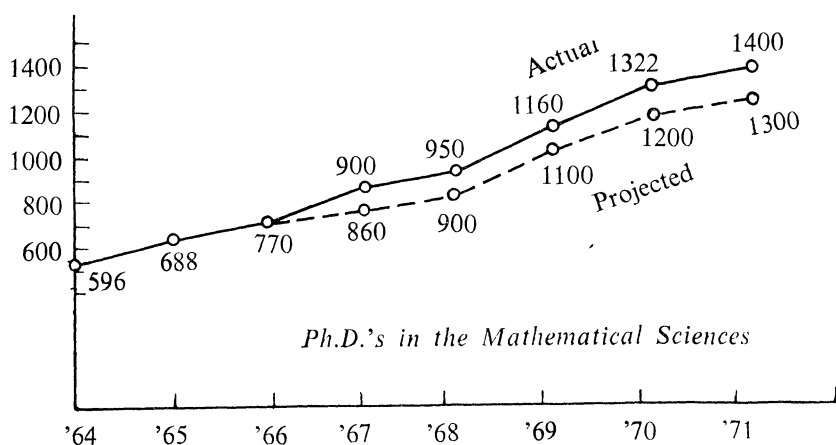


Fig. 2

Notice how different the situation is today from what was predicted. More Ph.D.'s and fewer B.A.'s. More teachers, and hence pressure for more mathematics classrooms, but fewer undergraduates interested in mathematics for mathematics sake. MORAL: If we value our jobs, it behooves us to see that our classrooms are filled with a greater concentration of students *whose first loyalties lie elsewhere*.

STRAW # 6. In a letter [8] to me in 1970, a loyal MAA member and MONTHLY reader expressed a concern about a disturbingly one-sided attitude towards mathematics which he detected in the pages of the MONTHLY. He said:

"Quite generally, recent issues of the MONTHLY seem to suggest feelings of self-complacency and smugness of some mathematicians which I can't believe are shared by most MAA members. Certainly, the exhibition of such feelings is not conducive to that attitude of increased social awareness and responsibility which in the opinion of some is demanded by the times. And by this I mean neither Viet-Nam nor Chicago, but rather, say the teaching of mathematics to non-mathematicians. Could it be that we need a little more insecurity?"

STRAW # 7. At the 1970 San Antonio Meeting, W. L. Duren [9] presented some slightly different views on the question "Are There Too Many Ph.D.'s in Mathematics?" During his talk he said:

"It is fair to say that, in the Federal planning which directed the support for an expanded graduate program in mathematics, there never was any indication that what was needed was more Ph.D. mathematicians who were traditionally trained to do research in some narrow field of pure mathematics. At best these research specialists were needed in greater numbers only as machine tools to produce more mathematicians. The real social needs for more mathematicians all came from the peripheral aspects of the field: for the computer revolution, aerospace efforts, optimization in engineering design, business management, for conceptual models to push forward in life and social sciences, and for more teaching to help young people to get jobs in a technical world. Besides this, an expanded graduate program in mathematics was needed to support the production of more Ph.D.'s in engineering and physical sciences."

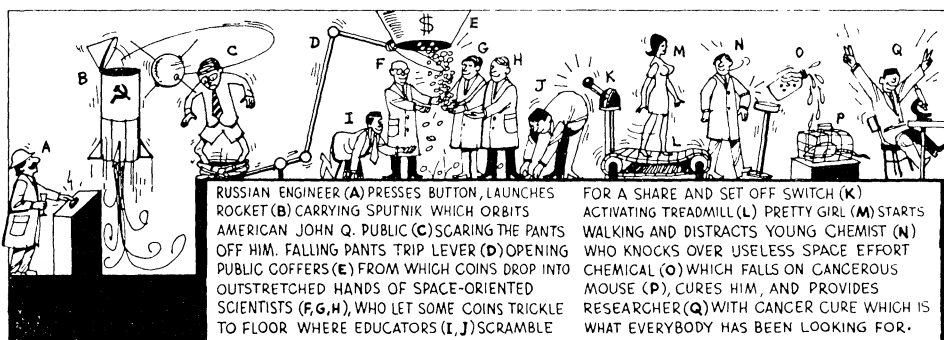
STRAW # 8. The February 7, 1971 issue of the Washington Post contained an article by Daniel S. Greenberg [10] entitled "Prestigious Science has Feet of Clay." It was a clear statement of the frustration which the "other" world feels with a science community which seems always to consume large amounts of money and prestige, pour forth large quantities of advice, but not to relieve in any systematic and regular way the awful suffering of Mankind. The article begins:

"The Republic has no need of scientists," declared the president of the French revolutionary tribunal as he sentenced the chemist Lavoisier to the guillotine. The wish that it were so may have strayed across the minds of some present-day officials, including Messrs. Nixon, Kosygin and Brezhnev and perhaps even Mao, as each in his own setting has confronted that peculiar and restless conglomeration known as "the scientific community."

"Seemingly insatiable for public funds, prickly and righteous when the public seeks a say about the use of those funds but frequently unhesitant to pronounce on political and social

issues, scientists constitute a group that is at once indispensable and often indigestible. Robert M. Hutchins encountered the tribe in his long ago days as head of the University of Chicago, and later concluded: 'A scientist has a limited education. He labors on the topic of his dissertation, wins the Nobel Prize by the time he is 35 and suddenly has nothing to do... He has no alternative but to spend the rest of his life making a nuisance of himself.'

For a touch of humor with a sharp edge, Greenberg's article is printed next to the following cartoon after the late Rube Goldberg.



© 1970, American Chemical Society, Reprinted by permission from *Chemical and Engineering News*, Vol. 48, December 21, 1970, page 5.

STRAW # 9. A recent issue of the NOTICES of the American Mathematical Society contained a "Letter to the Editor" from Mary B. Williams. Her letter reads, in part, [11]:

"The thing that an interdisciplinary mathematician most needs is an understanding of the relationship of mathematics to the real world. The best mathematics comes from a deep intuitive understanding of the structure of some portion of the real world; pure mathematics is useful because it starts from mathematical structures which were abstracted from the structure of the real world by earlier, non-pure mathematicians. But at present mathematicians are taught that mathematics is a free creation of the human mind, justified not by its connection to the real world but by its own intrinsic beauty. I realize that this philosophy solves (or, rather, relegates to the nether regions) some extremely difficult problems concerning the connection of mathematics with reality, but it is false. (Possibly most first-rate pure mathematicians would agree that it is false; nevertheless, it pervades their teaching.) Mathematicians expect usefulness to be an inevitable by-product of mathematical beauty; they justify this expectation by historical examples of pure mathematics which turned out to be useful, but they have no understanding of why these examples turned out to be useful and consequently when they want to do something useful they work on the assumption that mathematical beauty guarantees ultimate usefulness. This leads them to feel that a superficial understanding of the real world problem, together with their own mathematical creativity, is all that is necessary to do worthwhile work; and it leads them to reject as irrelevant the objection that their results don't solve any problem the scientist is interested in. Naturally the science departments do not want to hire mathematicians with this attitude."

STRAW # 10. At the 1971 MAA Summer Meeting in Laramie, Lowell Paige [12] gave a talk entitled "Public Understanding of Science and Its Implications

for Mathematics.” Every thoughtful member of the Association should read his talk which is reprinted in the MONTHLY, February 1971. Toward the end of his remarks, Dean Paige made the following comments:

“Let us look at some of the consequences which the mathematical community must face. First, it is to be expected that a major portion of any additional funds recommended for the National Science Foundation this year will be assigned to interdisciplinary programs directed at the problems of society. Even a casual reading of the testimony before the Special Subcommittee on NSF of the Senate reveals this fact. Hence, the fiscal support available for fundamental research in science, including the mathematical sciences, will remain approximately the same as last year....

“The most widely discussed criteria proposed for the assignment of priorities to scientific research are those advanced by Dr. A. Weinberg of the Oak Ridge National Laboratories. The criteria of justification proposed for the support of science were: technological merit, scientific merit, and social merit. It is in the discussion of scientific merit that he states, ‘I would therefore sharpen the criterion of scientific merit by proposing that, other things being equal, that field has the most scientific merit which contributes most heavily to and illuminates most brightly its neighboring scientific disciplines.’

“If the preceding is taken without modification as a reasonable basis for the allocation of funds within the National Science Foundation, then the mathematical sciences section will need all of the assistance the professional societies can provide for the justification of their requests. To illustrate my concern, I note that in Weinberg’s discussion of scientific merit preceding his recommendations he appeals to the following comment of von Neumann: ‘As a mathematical discipline travels far from its empirical source, or still more, if it is a second or third generation only indirectly inspired by ideas coming from reality, it is beset with grave danger. It becomes more and more pure aestheticizing, more and more purely *l’art pour l’art*. This need not be bad if the field is surrounded by correlated subjects which still have closer empirical connections or if the discipline is under the influence of men with an exceptionally well developed taste. But there is grave danger that the subject will develop along the line of least resistance, that the stream, so far from its source, will separate into a multitude of insignificant branches, and that the discipline will become a disorganized mass of details and complexities.’

“To be brief, the appeal is to relevance; not in the sense attached to relevance by students but in the intellectual context of unifying concepts. I do not interpret von Neumann’s remarks to be a clarion call for slavish devotion to the applications of mathematics; but I am certain that Weinberg and other scientists are not so inclined.

“Many mathematicians have expressed the need for our courses and research efforts to reflect the relation between various areas of mathematics as well as to the applications to other disciplines.

“I would propose that our writing include more than a feeble pass at articles designed to illustrate the unifying aspects of abstract concepts for the non-mathematical scientist. The initial effort of the COSRIMS reports must be continued if we are to convince our scientific colleagues that the plea in Hardy’s toast, “Here’s to pure mathematics. May it never have any use,” has not been fulfilled. Thus, I find articles of the nature of Saunders MacLane’s in the June/July issue of the MONTHLY to be of considerable importance....

“I have devoted considerable time to what might appear to be the selfish interests of faculty members. Now I wish to consider the important component of our concern: the students. What will be the effect of present attitudes upon our students?

“There is no doubt in my mind that the growing contention that science and technology are insensitive to our social problems is driving undergraduates from Science and Mathematics

to the Social Sciences. This can only result in further alienation from mathematics and I submit that one of our curriculum disaster areas is in courses designed for non-mathematics majors in addition to those service courses we provide for the various disciplines.

"Even if we choose to ignore the nonmathematics majors, our undergraduate majors cannot help but notice the reduction in fellowships and research assistantships for graduate study. It is estimated that the reduction will be approximately 20% this year. Is it any wonder that students are discouraged when the prospects for support are diminishing? And to this distressing note, we might add the publicity of an oversupply of Ph.D.'s which has been widely discussed in the mathematical community.

STRAW \neq 11. This last straw is anecdotal and also anonymous, because for all I know the story is still being acted out. A friend, chairman of the mathematics department at a major university, recently told me that the chemistry department in his university is currently debating a proposal to eliminate the requirement that its majors take a sophomore course given by the mathematics department. The proposal is to require instead that they take a mathematics course offered by the chemical engineering department. The mathematics department, they say, does not teach the mathematics they wish their majors to have and the chemical engineering department does. "I can hardly criticize the quality of the mathematics course the chemical engineers offer or their competence to offer it," says my friend. "After all the man in charge of it was formerly the chairman of the mathematics department at X University." X is also the name of a major and respected university.

When I heard this story I instantly recalled a statement made to me years ago by the Head of a Civil Engineering Department: "The Mathematics Department has us over a barrel, right now," he said, "but we are stockpiling mathematicians in our department for a future rainy day."

Have you taken a look at the weather reports lately?

3. Where does the wind blow. The past quarter-century has seen unprecedented growth in mathematics, particularly in the United States. On the threshold of the 70's the discipline stands as a strong force in the world of science and ideas. This world has accepted willingly — almost eagerly — our pronouncements about the importance and power of mathematics in a scientific age. It has even accepted, albeit somewhat less eagerly, our claims about the beauty of mathematics. The widening sphere of influence of mathematics in science has been noted by the world and the world has supported us in a style which befits our stature.

However, there has been a sudden subtle shift in the wind. Dispatches bring news of unrest in some of the farflung colonies. The treasure ships have been returning to port riding somewhat high lately. The flow of recruits from the hinterlands has fallen off. The world has begun to remind us that when we claim that a discipline is *basic central* and *powerful*, we are referring to its impact outside of itself. The world grumbles that it cannot afford, at this moment in history, the luxury of elegance for elegance's sake. It hardly has time to play chess any more, sorry, good game and all that, but no time any more.

The change in attitude of the world may well be very good for mathematics. Our discipline has always grown in cycles, which bring it closer at some times than others to the influences of the real world. Even in times of great abstraction and purity, mathematics drinks from the wells of science and society. But like any living organism, mathematics does have the option of remaining away from the well. Eton doesn't have to change. It is only the consequences of such a decision which are inevitable. Suicide is only fatal, it is not unavailable.

The future of mathematics in the next decade or two is not in our hands, yours and mine. It is in the hands of the students who populate our undergraduate classrooms right now. Mathematics has *no* future unless they stay there and emerge at the end of four years with a desire for mathematics as a career. And today's youth will not be bought by promises of elegance. They want to know what mathematics has to do with the price of eggs.

On a more mundane level, but one which we certainly cannot ignore, the students in our undergraduate classrooms pay our salaries. It seems quite possible that fewer of these students will be mathematics majors in the immediate future, and it is a matter of hard and selfish economic necessity for us to pay more attention than we have to what we are saying to these students who are more interested in the price of eggs than in mathematics.

This means a significant change in attitude in the teaching of undergraduate mathematics toward questions of the relationship of mathematics to other concerns. It is not just a matter of teaching more applied mathematics, or even of putting more examples and applications into our regular courses — although it does involve this. It is a matter of viewing our subject as a part of the larger efforts of man to describe, understand and control his physical, social and intellectual environment.

What does this mean for the average teacher in the average classroom in the the average department of mathematics? I cannot give any tested recipes for relevance. I don't believe that there are any. We are all individuals with individual tastes in mathematics. We adhere to different special fields within mathematics and these differences filter down even to our undergraduate classrooms. Relevance means one thing to a specialist in partial differential equations, another thing to a number theorist. Nevertheless, a few general observations might be made which apply to nearly all of us.

1. Just turn over in your mind each day the fact that no part of mathematics is totally and inherently immune to applicability and totally divorced from inspiration from outside of mathematics. Be sensitive to and interested in any little bridge between mathematics and other concerns that you happen upon in your travels, even though it is a tiny, seldom traveled footbridge. Even if it is only a narrow plank it may be used to cross the interface and there may be an idea on the other side which sheds new light for you on mathematics. If you are sensitive to these little

bridges, even though you have no need to cross any of them, that sensitivity will be passed on to your students, many of whom won't accept mathematics without it.

After I had given an earlier version of this talk, a young instructor came up to me and asked what he could do to make mathematics more relevant in his classroom. He is a number theorist, and he was in the fortunate position to spend most of his time teaching number theory to undergraduates. "How do you make number theory into applied mathematics?" he asked ruefully. I hadn't thought about my answer to such a question in advance, so I could only stammer out some such general pep-talk as I have made immediately above. Thinking of it later, I wished I had remembered a beautiful bit of applied mathematics I had come across years ago during one of my rare experiences in teaching number theory to an undergraduate class. The example is in Ore's charming book, "Number Theory and its History." Indices modulo an integer are used to obtain a simple set of rules for splicing telephone cables in such a way as to minimize interference, or "cross-talk", between circuits. Encountered unexpectedly in the middle of a course in the purest of pure mathematics, the application is a refreshing breath of air from outside. Of no particular importance by itself, it does shine a new light on the mathematics.

2. When you serve on a text-book selection committee insist that one of the elements to be considered in evaluating a given book is the picture it presents of mathematics in a broad intellectual context. How well does the book handle the available ties between that particular mathematics and other concerns, both for input and for output, motivation and application. Clearly, this isn't the only criterion for choosing a book and in many cases it may be a minor one, but don't leave it out of the equation.

3. In choosing topics for undergraduate seminars and colloquia, choose subjects which involve cross-fertilization between mathematics and other fields. This may take you far from your own field of competence in mathematics, but it is refreshing sometimes to sit in a seminar in which you know very little more than the students about the mathematical subject under discussion. One of the masters in finding intriguing ties between very pure mathematics and an amazing collection of other fields is Victor Klee. I urge you all to show his two films, "Shapes of the Future, I and II" to your undergraduates. He can find fascinating connections between unsolved problems in geometry and combinatorics and problems of current interest in solid state physics, virology, organic chemistry, botany and other fields. A number of these are reported in the Research Problems section of the MONTHLY, a rich source of mathematical topics with frequently unexpected applications to other fields.

4. Be open and sensitive to situations in which inter-disciplinary courses or seminars would be feasible and interesting on your campus. Our habits of disciplinary thinking are so strong, that inter-disciplinary courses are not easy to establish and

are often difficult to sustain. In the presence of students, our worst chauvinistic instincts have a way of coming to the surface. However, as a one-time participant in an inter-disciplinary course which lasted more than ten years, I have come to believe that no matter how brief the marriage, both partners (to say nothing of the students) reap rich benefits.

5. Establish cordial relations with individuals in other departments. Attend their seminars and colloquia occasionally. Suggest subjects for their seminars or colloquia when you know of topics in their fields where significant mathematization has occurred. Invite them to attend your seminars and colloquia when speakers or topics appear which might interest them even peripherally.

6. Ask yourself ten times each week, whether the cross-fertilization between mathematics and other fields, is not a subject of potential interest for any mathematician regardless of his mathematical interest. It may not become a central part of his mathematical life, but neither are many other things he finds intellectually interesting. For that matter, the biologist and even the physicist have not traditionally found the mathematicians' first love all that interesting. But is it not possible — not probable, perhaps, but just possible — that the reason for the lack of interest from our scientific colleague, is that no mathematician has ever been willing to talk with him, in terms comprehensible to the intelligent layman, about what he is doing?

I have put your patience to a severe test in dragging you through parable and metaphor. I ask you just once more to give your imagination free rein, as we make one more stop at the game of chess. Suppose someone were to discover that the rules and strategies of chess lead, on proper interpretation, to a strategy for traffic control which solved forever the problem of optimizing the flow of traffic in a large and congested urban area (the chess board) so that large masses of vehicles could move from one area to another, with virtually no accidents and absolutely minimal delays. What would this do for chess? (It is obvious what it would do for modern urban society.) Bobby Fischer and Boris Spassky would probably give scant recognition to the new applied chess. In fact, they and other chess purists would probably treat the development with disdain and even alarm. Traffic control is, after all, far from what chess is all about. But I wonder if the publicity alone would not cause a boom of unheard of proportions in chess playing. I wonder if sales of chess sets wouldn't double over a period of years, support of chess clubs triple, grants in basic research in chess quadruple. It is not unlikely that the world's champion chess player of the 1980's might be attracted to the game in 1971 because of this startlingly new bridge between the real world and the red, white and ivory world of the chessboard. Who would come off better in this game, the traffic engineers or the chess players?

I have tried to present some of the background and reasons for a conviction which is growing in my mind that some of the most significant events in mathematics in the 70's may well occur in the undergraduate classrooms of our nation. Not only do we begin the process of procreation there, we also implant indelible impressions

of the nature and value of mathematics to a significant segment of the world. If mathematics stands in the world like England did at the turn of the century, then we — you and I — are presiding over the Etons of Mathematics. If we lose contact with the essential ferment which is going on out there, then the world may simply walk away, not only from us but from mathematics. It is a challenge, a very real challenge, and it is one reason why I am proud to be associated with an organization, the MAA, which was created and exists exactly and exclusively to help us meet that challenge in the classroom. I hope that you will share my pride and that together we can insure the strength of mathematics for tomorrow by making it vital, meaningful, and (please forgive one last use of the overworked word) relevant for our students today.

Based on a talk presented at the meeting of several Sections of the MAA during 1971–72.

References

1. AMERICAN MATHEMATICAL MONTHLY, Volume 77, No. 6, June/July, 1970, p. 623.
2. AMERICAN MATHEMATICAL MONTHLY, Volume 77, No. 5, May 1970, pp. 514–515.
3. Internal C. U. P. M. document.
4. AMERICAN MATHEMATICAL MONTHLY, Volume 77, No. 6, June/July, 1970, pp. 626–641.
5. SCIENCE, Volume 167, 9 Jan., 1970, p. 141. (Reprinted with permission of the publisher, The American Association for the Advancement of Science.)
6. The Washington Post, Sunday, April 26, 1970, p. B5. (Reprinted with permission.)
7. NOTICES of the American Mathematical Society, Volume 19, No. 2, p. 119 © 1970. (Reprinted with permission of the American Mathematical Society.)
8. Letter to A. B. Willcox from Peter Henrici.
9. AMERICAN MATHEMATICAL MONTHLY, Volume 77, No. 7, June/July, 1970, pp. 641–646.
10. The Washington Post, Sunday, February 7, 1971, p. C1. (Reprinted with permission.)
11. NOTICES of the American Mathematical Society, Volume 18, No. 3, pp. 502–503, © 1971. (Reprinted with the permission of the publisher, The American Mathematical Society.)
12. AMERICAN MATHEMATICAL MONTHLY, Volume 78, No. 2, February 1971, pp. 130–142.

MATHEMATICAL NOTES

EDITED BY ROBERT GILMER

The present backlog for this Department is substantial. Until further notice, new manuscripts cannot be accepted. This moratorium will probably continue until June 1, 1973; authors are requested to hold their manuscripts pending a further announcement.

AN IDENTITY SATISFIED BY DERIVATIONS OF A PURELY INSEPARABLE FIELD

F. P. CALLAHAN, Pennsylvania State University

Introduction. Let k be a field such that $\text{char } k \neq 0$ and let $k(\alpha)$ be the purely inseparable extension of k obtained by adjoining to k an α which satisfies the ir-

reducible equation $\alpha^p = a$, where a is in k . Let ϕ be a k -linear derivation of $k(\alpha)$ (so that $\phi(cx) = c\phi(x)$ or, equivalently, $\phi(c) = 0$, for c in k and x in $k(\alpha)$). Then it is easy to see that if $\phi \neq 0$ there exists a unique u in $k(\alpha)$ such that $\phi = (1/u)D$, where D (which may appropriately be called $d/d\alpha$) is the k -linear derivation of $k(\alpha)$ for which $D(\alpha) = 1$. The result proven in the paper can now be stated:

THEOREM. *Let $\phi (\neq 0)$ be a k -linear derivation of $k(\alpha)$ so that there exists a unique u in $k(\alpha)$ for which $\phi = (1/u)D$, where $D = d/d\alpha$. Then ϕ satisfies the identity:*

$$\phi^p = b\phi, \text{ where } b = -D^{p-1}(u)/u^p.$$

Proof of Theorem. The general idea of the proof is to view ϕ as a k -linear endomorphism of the underlying vector space of the algebra $k(\alpha)$.

As a first step in the proof we remark that $(1, x, x^2, \dots, x^{p-1})$ is a basis for $k(\alpha)$ over k if and only if x is not in k . This is easily shown. Next we consider two cases: in the first ϕ is nilpotent and in the second it is not.

Case 1. $\phi^p = 0$. In this case, by considering $\phi(\alpha), \phi^2(\alpha), \dots, \phi^{p-1}(\alpha)$, we find an n such that $\phi^{n-1}(\alpha) \neq 0$ but $\phi^n(\alpha) = 0$. If $\phi^{n-1}(\alpha) = c$, then c must be in k , because otherwise the remark in the paragraph above shows that ϕ is identically zero. Thus, if y is defined to be $\phi^{n-2}(\alpha/c)$, then $\phi(y) = 1$ and ϕ is d/dy ; also, $\phi = (1/u)D$ where $u = Dy$. Since $D^p = 0$, as is easily seen, this gives $D^{p-1}u = 0$ so that the b defined in the statement of the theorem is zero in this case and the theorem is seen to be true in Case 1.

Case 2. $\phi^p \neq 0$. In this case it is evident that ϕ must have at least one non-zero eigenvalue; let such an eigenvalue be λ and if it is not already in the groundfield k adjoin it to k to obtain a larger groundfield k' and algebra $k'(\alpha)$. The extension of ϕ to a k' -linear derivation of $k'(\alpha)$ presents no problems, and we continue to call it ϕ . Now let x be an eigenvector of ϕ such that $\phi(x) = \lambda x$. Since ϕ is a derivation this implies that $\phi(x^i) = ix^{i-1}\phi(x) = i\lambda x^i$, so that $(1, x, x^2, \dots, x^{p-1})$ is a complete set of eigenvectors for ϕ , the corresponding eigenvalues being $(0, \lambda, \dots, (p-1)\lambda)$. Since $i^p \equiv i \pmod{p}$, we see that $\phi^p(x^i) = b\phi(x^i)$ where $b = \lambda^{p-1}$. Since $(1, x, \dots, x^{p-1})$ is a basis this implies that $\phi^p = b\phi$. Also, since the eigenvalues are distinct it implies that $\theta^p = b\theta$ is the characteristic equation of ϕ and that it is the only equation of degree p satisfied by all the eigenvalues of ϕ .

To complete the proof it remains to show that $b = -(D^{p-1}(u))/u^p$. To do this, again let λ be an eigenvalue of ϕ and x the corresponding eigenvector. Define the k' -linear operator T by the formula $T = D + \lambda M(u)$ where $M(u)$ is the operation of multiplying by u (so that $M(u)(y) = uy = yu$). Define z_n by $z_n = T^n(1)$, where "1" is the identity of the algebra. Now an easy inductive argument shows that $D^n(x) = z_n x$.

Since $D^p = 0$ this implies that $z_p x = 0$; since $k'(\alpha)$ is a field and $x \neq 0$ this implies that $z_p = 0$. That is, $T^p(1) = 0$.

Now the definition of T implies that $T^p = T_p \lambda^p + T_{p-1} \lambda^{p-1} + \cdots + T_0$, where T_m is the sum of all monomial operators of the form $V_1 V_2 \cdots V_p$, each V being either $M(u)$ or D , and $M(u)$ occurring m times and D occurring $p - m$ times in each such monomial term. Thus the equation $T^p(1) = 0$ becomes $t_p \lambda^p + t_{p-1} \lambda^{p-1} + \cdots + t_0 = 0$, where $t_i = T_i(1)$.

Since λ is any eigenvalue of ϕ , we see that the equation above in λ can only differ from the characteristic equation, $\theta^p - b\theta = 0$, of ϕ , by a multiplicative constant. Thus, of the coefficients t_0, t_1, \dots, t_p , all but t_1 and t_p must vanish and we must have $b = -t_1/t_p$.

Now t_p and t_1 are easily computed as follows:

$$t_p = T_p(1) = (M(u))^p(1) = u^p$$

and

$$t_1 = (M(u)D^{p-1} + DM(u)D^{p-2} + \cdots + D^{p-1}M(u))(1) = 0 + 0 + \cdots + D^{p-1}(u).$$

Thus, $b = -D^{p-1}(u)/u^p$ and the proof is complete.

REMARKS. (1) The vanishing of all the t 's except t_1 and t_p provides some unobvious identities. An alternative proof of these identities can be given by means of the following identity which holds good in any commutative ring of characteristic p :

Let ϕ be a derivation of R and let u be a member of R . Let operator S be defined by $S(y) = \phi(y) + uy$, where uy is the product of u and y in R . Then $S^p(y) = \phi^p(y) + u^p y + vy$ where y is any element of R and $v = \phi^{p-1}(u)$.

This identity can be obtained from one given by Jacobson (*Lie Algebras*, Interscience Publishers, page 187, equation (63) with a and b replaced by D and $\lambda M(u)$, resp.).

(2) The referee has kindly pointed out that the equation $\phi^p = b\phi$ but not the explicit evaluation of b can also be obtained by way of Jacobson's Galois Theory for purely inseparable fields of exponent one. (See N. Jacobson, *Lectures in Abstract Algebra*, Van Nostrand, Princeton, N.J., Volume 3, page 190, equation 34.)

ON SUMS OF POWERS OF A NUMBER

VLADIMIR DROBOT, State University of New York at Buffalo

In this note we investigate some approximation properties of polynomials whose coefficients are 0, +1, or -1. The original problem, posed by M. Parnes [1] can be formulated as follows: Suppose we take a walk along the x -axis, starting from the origin, and at the time n we are allowed to take a step of length 0, $+t^n$, or $-t^n$ (t is fixed). Which points can we approach as closely as we wish? If $0 < t < 1$, it is clear that we cannot get outside the interval $[-(1-t)^{-1}, +(1-t)^{-1}]$. Since $\sum_{n=N}^{\infty} t^n \geq t^{N-1}$ if $\frac{1}{2} \leq t < 1$, it is easy to prove that for such t every point in the

above interval can be approximated as closely as we wish. If $t > 2$ and if $\varepsilon_n = 0$, $+1$, or -1 , $\varepsilon_N \neq 0$ then

$$|\varepsilon_0 + \varepsilon_1 t + \cdots + \varepsilon_N t^N| \geq t^N - (1 + t + \cdots + t^{N-1}) > t^N(1 - (t-1)^{-1})$$

which tends to ∞ as $N \rightarrow \infty$. From now on we shall restrict our attention to the case $1 < t < 2$. Let \mathcal{P} be the set of polynomials with coefficients 0, $+1$, or -1 and let $\mathcal{P}(t) = \{p(t) : p \in \mathcal{P}\}$. The above remarks show that if $0 < t < 1$ then $\mathcal{P}(t) \subseteq [-(1-t)^{-1}, (1-t)^{-1}]$ and if $t \geq 2$ then $\mathcal{P}(t)$ is discrete. For $1 < t < 2$, $\mathcal{P}(t)$ is dense in the line for all but (possibly) a countable number of t 's as the following result shows.

THEOREM. *If $1 < t < 2$ is not a root of any of the polynomials in \mathcal{P} then $\mathcal{P}(t)$ is dense in the line.*

Proof. First of all it is enough to prove that 0 is a cluster point of the set $\mathcal{P}(t)$. Indeed assume that this is so. For any two positive integers n and k there can be found an integer $m \geq n$ and a polynomial $p \in \mathcal{P}$ such that

$$(1) \quad t^{-k-1} \leq t^m p(t) < t^{-k}.$$

This is done by first choosing $p \in \mathcal{P}$ so that $0 < p(t) < t^{-k-n}$, and then taking $m \geq n$ such that $t^{-k-m-1} \leq p(t) < t^{-k-m}$. It is clear that if p belongs to \mathcal{P} then so does $\pm x^m p(x)$. We construct now a sequence of polynomials $\{p_N\} \subseteq \mathcal{P}$ no two of which have terms in common and every one of which satisfies

$$(2) \quad t^{-k-1} \leq p_N(t) < t^{-k}.$$

This is done as follows. Let $p_1 \in \mathcal{P}$ be any polynomial in \mathcal{P} for which (2) holds. Assume the polynomials p_1, p_2, \dots, p_N are already chosen, that they have no terms in common, and they all satisfy (2). Let n be an integer larger than the degree of any of these polynomials and let $p_{N+1}(x) = x^n p(x)$ be a polynomial in \mathcal{P} satisfying (1). The sequence is hence defined inductively. It follows that $p_1 + p_2 + \cdots + p_N \in \mathcal{P}$ and

$$Nt^{-k-1} \leq p_1(t) + \cdots + p_N(t) \leq Nt^{-k}.$$

Since N and k are arbitrary and $t > 1$, we see that any number can be approximated as closely as we wish, by elements of $\mathcal{P}(t)$.

To prove that 0 is the closure of $\mathcal{P}(t)$ fix $\varepsilon > 0$, let \mathcal{P}_n be the set of polynomials in \mathcal{P} of degree at most n and let \mathcal{P}_n^+ be the set of polynomials in \mathcal{P}_n with non-negative coefficients. There are 2^{n+1} polynomials in \mathcal{P}_n^+ and if p is one of them then $p(t)$ lies between 0 and $1 + \cdots + t^n = (t^{n+1} - 1)(t-1)^{-1}$. If $p \neq q$ are in \mathcal{P}_n^+ then $p - q$ is in \mathcal{P}_n and $p(t) \neq q(t)$. Hence by the pigeon hole principle, there are two polynomials $p_n \neq q_n$ in \mathcal{P}_n^+ such that

$$(3) \quad 0 < |p_n(t) - q_n(t)| \leq 2^{-n}(t^{n+1} - 1)(t-1)^{-1}.$$

Since $1 < t < 2$, the right hand side of (2) is less than ε for large n .

The set $\mathcal{P}(t)$ is not dense, however, for all $1 < t < 2$. This was discovered by Prof. J. Isbell [2]. I am grateful to him for permission to include his example here as well as for some helpful conversations on the subject.

EXAMPLE (J. Isbell). If $t = \frac{1}{2}(\sqrt{5} + 1)$ is the golden mean then $\mathcal{P}(t)$ is discrete.

Proof. The golden mean t satisfies the equation $t^2 = t + 1$. We order all the polynomials $p \in \mathcal{P}$ lexicographically according to the decreasing exponents of the terms. More precisely, let

$$p(x) = a_0 + a_1 x + \cdots + a_N x^N, \quad q(x) = b_0 + b_1 x + \cdots + b_M x^M,$$

be two polynomials in \mathcal{P} . Let k be the largest integer for which $|a_k| \neq |b_k|$. We say that $p(x)$ is of lower rank than $g(x)$ if $a_k = 0$ and $b_k \neq 0$. If $|a_j| = |b_j|$ for all j , we arbitrarily say p is of lower rank than q if $p(\pi) < q(\pi)$. If $s \in \mathcal{P}(t)$ let $p_s(t)$ be the polynomial of lowest rank such that $s = p_s(t)$, let $h(s)$ be the degree of $p_s(t)$ and let $\alpha(s)$ be the last exponent before the first sign change. That is,

$$s = \pm (t^{h(s)} + \cdots + t^{\alpha(s)} - t^\beta \pm \cdots).$$

We claim that $\beta = \alpha(s) - 1$. Moreover, all the terms with exponents less than α are present in the above representation and their coefficients alternate in sign. Indeed, if $\beta < \alpha - 1$ then the terms $t^\alpha + 0t^{\alpha-1} - t^{\alpha-2}$ and $t^\alpha + 0t^{\alpha-1} + 0t^{\alpha-2}$ can be replaced respectively by $t^{\alpha-1}$ and $-t^{\alpha-1} - t^{\alpha-2}$ both of which are of lower lexicographical order. Also if $\gamma > 1$ and t^γ and $t^{\gamma-1}$ are present in $p_s(t)$ with opposite signs then $t^{\gamma-2}$ must also be present with the same sign as t^γ . Otherwise the terms $+ (t^\gamma - t^{\gamma-1} + 0t^{\gamma-2})$ and $\pm (t^\gamma - t^{\gamma-1} - t^{\gamma-2})$ can be replaced by the terms $\pm t^{\gamma-2}$ and 0, respectively, yielding a polynomial of lower rank. Finally then

$$\begin{aligned} s = p_s(t) &= \pm \left\{ t^h + \cdots + t^\alpha + \sum_{n=0}^{\alpha-1} (-1)^{\alpha-n} t^n \right\} \\ &= \pm \{ t^h + \cdots + t^\alpha - (t^\alpha \pm 1)(t + 1)^{-1} \}. \end{aligned}$$

It easily follows now that $|s| \rightarrow \infty$ as $h \rightarrow \infty$ and so $\mathcal{P}(t)$ is discrete.

References

1. M. Parnes, Problem # 15, Ridge Lea Problem Book, Mathematics Department, SUNY at Buffalo.
2. J. Isbell, Private communications.

A LOCAL MEAN VALUE THEOREM FOR ANALYTIC FUNCTIONS

ÅKE SAMUELSSON, University of Göteborg

The classical mean value theorem of differential calculus does not extend to the complex plane. The purpose of this note is to establish a local counterpart for analytic functions.

THEOREM. *If f is analytic in a domain containing z_0 then there is a neighborhood N of z_0 such that if z_1 is any point in this neighborhood then there exists a point z with*

$$\left| z - \frac{1}{2}(z_0 + z_1) \right| < \frac{1}{2} |z_1 - z_0|,$$

such that $f(z_1) - f(z_0) = (z_1 - z_0)f'(z)$.

A slightly weaker version of this theorem has been proved by J. M. Robertson [1]. As a matter of fact, with the additional assumption that $f''(z_0) \neq 0$, Robertson's proof yields our theorem.

Proof. We may assume that f has the form

$$f(z) = f(z_0) + (z - z_0)f'(z_0) + (z - z_0)^{k+1}h(z),$$

where $k \geq 1$ is an integer and $h(z_0) \neq 0$.

We may also assume, without loss of generality, that throughout the domain of analyticity we have

$$|h(z)| \geq \frac{1}{2} |h(z_0)| \quad \text{and} \quad |h'(z)| \leq 1.$$

It suffices to show that if the neighborhood $N = \{z; |z - z_0| < r\}$ is chosen so that $0 < r \leq |h(z_0)|/2(k+2)$ and $z_1 \in N$, then the function

$$f'(z) - \frac{f(z_1) - f(z_0)}{z_1 - z_0}$$

has exactly one zero in the domain

$$D = \left\{ z; \left| z - \frac{1}{2}(z_0 + z_1) \right| < \frac{1}{2} |z_1 - z_0|, \quad \left| \arg \frac{z - z_0}{z_1 - z_0} \right| < \frac{\pi}{k} \right\}.$$

A direct computation shows that

$$f'(z) - \frac{f(z_1) - f(z_0)}{z_1 - z_0} = \Phi(z) + h(z_1)\psi(z),$$

where $\Phi(z) = (z - z_0)^{k+1}h'(z) + (k+1)(z - z_0)^k(h(z) - h(z_1))$ and

$$\psi(z) = (k+1)(z - z_0)^k - (z_1 - z_0)^k.$$

If $z \in \partial D$, the boundary of D , then

$$\begin{aligned} |\Phi(z)| &\leq |z - z_0|^{k+1} |h'(z)| + (k+1) |z - z_0|^k \left| \int_{z_1}^z h'(\zeta) d\zeta \right| \\ &\leq (k+2) |z_1 - z_0|^{k+1}. \end{aligned}$$

If z is on the circular arc of ∂D , i.e., if $z = \frac{1}{2}(z_0 + z_1) + \frac{1}{2}(z_1 - z_0)e^{i2\theta}$, $|\theta| \leq \pi/k$, then

$$|\psi(z)|^2 / |z_1 - z_0|^{2k} = 1 + (k+1)((k+1)\cos^k\theta - 2\cos k\theta)\cos^k\theta.$$

Using the inequality

$$(k+1)\cos^k\theta - 2\cos k\theta \geq 0 \text{ for } |\theta| \leq \pi/k, \quad k = 1, 2, \dots,$$

readily established by induction, we see that $|\psi(z)| \geq |z_1 - z_0|^k$. If $k > 2$, then the boundary ∂D contains two line segments, namely $z = z_0 + t(z_1 - z_0)e^{\pm i\pi/k}$, $0 \leq t \leq \cos \pi/k$. On these line segments we have

$$|\psi(z)| = (1 + (k+1)t^k) |z_1 - z_0|^k \geq |z_1 - z_0|^k.$$

We have shown that $|\psi(z)| \geq |z_1 - z_0|^k$ on ∂D . Hence, for $z_1 \in N$ and $z \in \partial D$,

$$\left| \frac{\Phi(z)}{h(z_1)\psi(z)} \right| \leq \frac{k+2}{|h(z_1)|} |z_1 - z_0| < \frac{|h(z_0)|}{2|h(z_1)|} \leq 1.$$

By Rouché's theorem we conclude that the functions $\Phi + h(z_1)\psi$ and ψ have equally many zeros in D , namely one. This proves our theorem.

Reference

1. J. M. Robertson, A local mean value theorem for the complex plane, Proc. Edinburgh Math. Soc. (2) 16 (1968/69), 329-331.

A THEOREM ON SET INCLUSION IN METRIC SPACES

JAMES A. HEINEN, Marquette University, and ALBERT WILANSKY, Lehigh University

Let A and B be subsets of a metric space (X, d) . We shall show that under certain (essentially sharp) conditions, A will be contained in B if $\partial A \subset B$. This result has applications in the study of the stability properties of certain differential equations and to the variation of the spectrum of a Banach algebra element.

For any set A in a metric space (X, d) , let A' denote the complement of A , $C(A)$ the closure of A , and ∂A the boundary of A .

THEOREM 1. *Suppose A and B are relatively compact (i.e. $C(A)$ and $C(B)$ are compact) subsets of a non-compact metric space (X, d) with B' connected. Then the condition $\partial A \subset B$ implies $A \subset B$.*

Proof. By a theorem of Hausdorff (a proof is given in [1], Theorem 1) we can give X an equivalent unbounded metric d_1 . A and B are still relatively compact, and hence bounded, in (X, d_1) . Now assume that $\partial A \subset B$ and, for the purpose of contradiction, that there exists a point $x \in A$ such that $x \notin B$, i.e. such that $x \in B'$. Let $D_1 = C(A) \cap B'$ and $D_2 = C(A') \cap B'$. Clearly, $x \in D_1$, so that $D_1 \neq \emptyset$, the null set. Since A and B are both bounded and (X, d_1) is unbounded, it follows at once that $D_2 \neq \emptyset$. Furthermore,

$$D_1 \cup D_2 = [C(A) \cap B'] \cup [C(A') \cap B'] = B'.$$

$$\text{Now } C(D_1) = C[C(A) \cap B'] \subset C[C(A)] \cap C(B') = C(A) \cap C(B').$$

Hence

$$C(D_1) \cap D_2 \subset C(A) \cap C(B') \cap D_2 = C(A) \cap C(B') \cap C(A') \cap B' = \partial A \cap B'.$$

But since $\partial A \subset B$, ∂A and B' have no points in common, thus implying that $C(D_1) \cap D_2 \subset \partial A \cap B' = \emptyset$, and, in fact, that $C(D_1) \cap D_2 = \emptyset$. In a similar fashion it may be shown that $D_1 \cap C(D_2) = \emptyset$. Thus $B' = D_1 \cup D_2$ where $D_1 \neq \emptyset$, $D_2 \neq \emptyset$, and where $C(D_1) \cap D_2 = D_1 \cap C(D_2) = \emptyset$. That is to say, B' is the union of two non-void separated subsets. This contradicts the assumption that B' is connected. Hence there can exist no point $x \in A$ such that $x \notin B$, and thus $A \subset B$.

To show that each hypothesis of the theorem is required, consider the following cases in which $\partial A \subset B$ and yet $A \not\subset B$ (in each case d is the usual Euclidean metric):

(1) A not relatively compact. $X = R^2$, $A = \{x \in X: d(x, 0) \geq 1\}$, $B = \{x \in X: d(x, 0) \leq 2\}$.

(2) B not relatively compact. $X = R^2$, $A = \{x \in X: d(x, 0) \leq 2\}$, $B = \{x \in X: d(x, 0) \geq 1\}$.

(3) B' not connected. $X = R^2$, $A = \{x \in X: d(x, 0) \leq 2\}$, $B = \{x \in X: 1 \leq d(x, 0) \leq 3\}$.

(4) (X, d) not non-compact. $X = \{x \in R^2: d(x, 0) \leq 3\}$, $A = \{x \in X: d(x, 0) \leq 2\}$, $B = \{x \in X: 1 \leq d(x, 0) \leq 3\}$.

As indicated earlier, this result has applications in the study of the behavior of solutions of differential equations [2]. Consider the n -dimensional vector differential equation

$$(1) \quad \dot{x} = f(x, t),$$

where it is assumed that f is sufficiently smooth to guarantee unique solutions which depend continuously on initial data. Let $x(t; x_0, t_0)$ denote the (unique) solution of (1) satisfying $x(t_0; x_0, t_0) = x_0$. Under these conditions, $x(t; \cdot, t_0)$ is a homeomorphism from R^n to R^n . Since set boundaries and compactness are preserved under homeomorphisms, it can readily be shown, using Theorem 1, that if S_0 is a compact subset of R^n and S is a bounded subset of R^n with S' connected, then the condition $x(t; \partial S_0, t_0) \subset S$ implies the condition $x(t; S_0, t_0) \subset S$. This, of course, allows one to arrive at conclusions regarding the nature of solutions of equation (1) for

all $x_0 \in S_0$ by simply verifying these conditions for all $x_0 \in \partial S_0$. As might be expected, this is of great interest when studying the stability of solutions of equation (1).

Theorem 1 also leads to results in the study of Banach algebras. Let A be a Banach algebra with identity 1 and S a closed subalgebra with $1 \in S$. Then $\sigma(a, S) = \{z \in C: a - z1 \text{ has no inverse in } S\}$ is, for each $a \in A$, a compact subset of the complex plane C ([3], p. 261, Theorem 3). Let $\rho(a, S)$ be the complement of $\sigma(a, S)$.

THEOREM 2. *If $\rho(a, A)$ is connected, $\sigma(a, S) = \sigma(a, A)$; thus $\sigma(a, S)$ is independent of S .*

Proof. “ \supset ” is trivial. To prove “ \subset ”, we note ([3], p. 266, Problem 23) that any boundary point of $\sigma(a, S)$ is in $\sigma(a, A)$ so Theorem 1 applies.

COROLLARY. *Suppose that for a certain $a \in A$ there exists S_1 such that $\sigma(a, S_1)$ is real (or more generally, is nowhere dense and has connected complement); then $\sigma(a, S)$ is independent of S .*

References

1. V. L. Klee, Jr., Some characterizations of compactness, this MONTHLY, 58 (1951) 389–393.
2. J. A. Heinen, Set Stability of Dynamical Systems, Ph. D. Dissertation, Marquette University, Milwaukee, Wisconsin, 1969.
3. A. Wilansky, Functional Analysis, Ginn-Blaisdell, Waltham, Mass., 1964.

CIRCLE GROUPS OF NILPOTENT RINGS

J. C. AULT AND J. F. WATTERS, The University of Leicester, England

A radical ring R is equal to its Jacobson radical and is therefore a group under the \circ -operation given by

$$a \circ b = a + b + a \times b,$$

where $+$ and \times denote the addition and multiplication in R . The group thus formed is called the **circle group** (Kruse, [3]). If R is a nilpotent ring of index n , say, that is $R^n = 0$ (any product of n elements is zero) but $R^{n-1} \neq 0$, then R is a radical ring and its circle group is a nilpotent group of class at most $n - 1$, as we show in the remark below. It is of interest to know which nilpotent groups arise as the circle groups of nilpotent rings. Kruse [3] has given necessary conditions for a finite nilpotent group to be a circle group and from these it can be deduced that not every nilpotent group of class 3 is a circle group. On the other hand, every Abelian group (nilpotent of class 1) is a circle group (of a zero ring, that is nilpotent of index 2, in fact, but also in many cases as circle groups of nilpotent rings of index greater than 2). The purpose of the present note is to consider the case of nilpotent groups of class 2. Kaloujnine [2] has already established that a large class of such groups are circle groups, but our method is more general and deals with all finite groups as well as some infinite groups.

REMARK. To show that a nilpotent ring R of index n has a circle group which is nilpotent of class at most $n - 1$, we consider the chain

$$(1) \quad R \supset R^2 \supset \cdots \supset R^{n-1} \supset R^n = 0,$$

as a series of subgroups of the circle group of R .

If $x \in R^k$, where $1 \leq k < n$, $y \in R$ and $y' \in R$ is such that $y' \circ y = 0$, then

$$y' \circ x \circ y = x + y' \times x + x \times y + y' \times x \times y$$

belongs to R^k . Hence R^k is a normal subgroup of the circle group of R . Furthermore, if $x' \in R$ is such that $x' \circ x = 0$, then

$$x' \circ y' \circ x \circ y \in R^{k+1}$$

so that (1) is a central series in the circle group of R , which is therefore a nilpotent group of class at most $n - 1$.

Let G be a nilpotent group of class 2, that is the centre Z of G is such that the factor group G/Z is Abelian. We recall here that, in such a group, we have the commutator identities

$$[ab, c] = [a, c][b, c]$$

and

$$[a, bc] = [a, b][a, c],$$

where $[a, b] = a^{-1}b^{-1}ab$ and a, b and c are elements of G . These identities will be used in the subsequent calculations without further reference.

We begin by establishing a necessary and sufficient condition for G to be the circle group of a nilpotent ring of index 3.

THEOREM 1. *The nilpotent group G of class 2 is the circle group of a nilpotent ring of index 3 if and only if there is a mapping m from the Cartesian product $G \times G$ into Z such that for all g, h and k in G ,*

- (i) $m(gh, k) = m(g, k)m(h, k)$,
- (ii) $m(g, hk) = m(g, h)m(g, k)$,
- (iii) $m(m(g, h), k) = m(g, m(h, k)) = e$,
where e denotes the identity element in G , and
- (iv) $m(g, h)\{m(h, g)\}^{-1} = [g, h]$.

Proof. If G is the circle group of a ring R , with $R^3 = 0$, then it is not difficult to verify that the mapping m given by

$$m(g, h) = g \times h$$

takes its values in the centre of G and satisfies the conditions (i) to (iv).

Conversely, if G is a nilpotent group of class 2 and there is a mapping m satisfying the conditions (i) to (iv), then we define

$$g + h = hg m(g, h) \quad \text{and} \quad g \times h = m(g, h)$$

for all g and h in G . It is straightforward to verify that these definitions make G into a ring R which is nilpotent of index 3. The details are left to the reader.

We mention that condition (iii) implies that

$$m(g, e) = e = m(e, g)$$

for all g in G . It then follows that the element e of G is the zero element of the ring. The negative of an element g of G may be expressed in the form

$$-g = g^{-1}\{m(g, g^{-1})\}^{-1}.$$

REMARK. Condition (iv) implies that $R^2 \neq 0$. It is worth noting that, in the verification of the ring axioms, condition (iv) is needed only for the commutativity of the addition, but it is the one which is the most significant. There are many mappings which satisfy (i), (ii) and (iii), (for example, $m(g, h) = e$ or $m(g, h) = [g, h]$), but it is more difficult to find ones which also satisfy (iv). In the next theorem we show how to construct such a mapping m when G is finite.

THEOREM 2. *If G is a finite nilpotent group of class 2, then G is the circle group of a nilpotent ring of index 3.*

Proof. Since G is finite, the Abelian group G/Z is isomorphic to a direct product of a finite number of finite cyclic groups. Thus we can choose a set of independent generators Za_1, Za_2, \dots, Za_r for G/Z having orders n_1, n_2, \dots, n_r , say. Then, given any g in G , there are unique exponents $\alpha_1(g), \alpha_2(g), \dots, \alpha_r(g)$ such that

$$\begin{aligned} Zg &= (Za_1)^{\alpha_1(g)}(Za_2)^{\alpha_2(g)} \dots (Za_r)^{\alpha_r(g)} \\ &= Za_1^{\alpha_1(g)}a_2^{\alpha_2(g)} \dots a_r^{\alpha_r(g)} \end{aligned}$$

and $0 \leq \alpha_i(g) < n_i$ for $i = 1, 2, \dots, r$. Put

$$(2) \quad m(g, h) = \prod_{i < j} [a_i, a_j]^{\alpha_i(g)\alpha_j(h)},$$

where $i, j = 1, 2, \dots, r$. Since $[a_i, a_j]$ is in Z and the exponents are uniquely determined, this does define a mapping from $G \times G$ into Z . It remains to check conditions (i) to (iv) of Theorem 1.

(i) It is necessary to calculate the exponent $\alpha_i(gh)$. Now

$$Zgh = (Zg)(Zh) = Za_1^{\alpha_1(g) + \alpha_1(h)}a_2^{\alpha_2(g) + \alpha_2(h)} \dots a_r^{\alpha_r(g) + \alpha_r(h)}$$

so that

$$(3) \quad \alpha_i(gh) \equiv \alpha_i(g) + \alpha_i(h) \pmod{n_i}.$$

Since $[a_i, a_j]^{n_i} = [a_i^{n_i}, a_j] = e$ we have

$$[a_i, a_j]^{\alpha_i(gh)\alpha_j(k)} = [a_i, a_j]^{\alpha_i(g)\alpha_j(k)}[a_i, a_j]^{\alpha_i(h)\alpha_j(k)}.$$

Therefore

$$\begin{aligned} m(gh, k) &= \prod_{i < j} [a_i, a_j]^{\alpha_i(gh)\alpha_j(k)} \\ &= \prod_{i < j} [a_i, a_j]^{\alpha_i(g)\alpha_j(k)} \prod_{i < j} [a_i, a_j]^{\alpha_i(h)\alpha_j(k)} \\ &= m(g, k)m(h, k). \end{aligned}$$

(ii) This condition follows in the same way as (i).

(iii) If z is in Z , then $\alpha_i(z) = 0$ for all $i = 1, 2, \dots, r$.

Hence

$$m(g, z) = m(z, k) = e$$

for all z in Z and g and k in G . Condition (iii) follows since m takes values in Z .

(iv) Suppose

$$g = z_1 a_1^{\alpha_1(g)} a_2^{\alpha_2(g)} \dots a_r^{\alpha_r(g)}$$

and

$$h = z_2 a_1^{\alpha_1(h)} a_2^{\alpha_2(h)} \dots a_r^{\alpha_r(h)},$$

where z_1 and z_2 are in Z . Then

$$\begin{aligned} [g, h] &= \prod_{i, j=1}^r [a_i, a_j]^{\alpha_i(g)\alpha_j(h)} \\ &= \prod_{i < j} [a_i, a_j]^{\alpha_i(g)\alpha_j(h) - \alpha_i(h)\alpha_j(g)} \end{aligned}$$

where $i, j = 1, 2, \dots, r$. Thus

$$\begin{aligned} m(g, h) \{m(h, g)\}^{-1} &= \prod_{i < j} [a_i, a_j]^{\alpha_i(g)\alpha_j(h)} \prod_{i < j} [a_i, a_j]^{-\alpha_i(h)\alpha_j(g)} \\ &= [g, h]. \end{aligned}$$

This completes the proof of the theorem.

Can every infinite nilpotent group of class 2 occur as the circle group of a nilpotent ring? It is our conjecture that every nilpotent group G (centre Z) of class 2 is the circle group of a nilpotent ring of index 3, but have not yet been able to prove this in general. However, we can show that this conjecture is true in the following cases.

Case 1. The group G/Z is a direct product of cyclic groups.

The formal definition of the function m is exactly the same as in (2). In the case when G/Z has an infinite factor with generator Za_i , say, then $\alpha_i(g)$ can be any integer and the congruence (3) becomes

$$\alpha_i(gh) = \alpha_i(g) + \alpha_i(h).$$

In the case when G/Z has infinitely many factors, these factors have to be indexed

by a well-ordered index set. Since, for a given element g of G , only finitely many of the exponents $\alpha_i(g)$ will be non-zero, there will only be finitely many non-identity factors in the right-hand side of (2) and so m is well-defined.

Case 2. The group G/Z is a torsion group.

This case is more difficult but may be reduced to the previous one by first decomposing G/Z into its p -components and then considering, in each of these components, a basic subgroup, which by definition is a direct product of cyclic groups (Fuchs [1, p. 98]).

Case 3. Every element of Z has a unique square root.

Here we set $m(g, h) = [g, h]^{\frac{1}{2}}$ and it is not difficult to verify that this satisfies conditions (i) to (iv). The ring so obtained is essentially the same as the one discussed by Kaloujnine [2].

This case includes the case when Z has odd exponent.

Whether the conjecture is true in general remains an open question.

References

1. L. Fuchs, *Abelian Groups*, Pergamon, London, 1960.
2. L. Kaloujnine, Zum Problem der Klassifikation der endlichen metabelschen p -Gruppen, *Wiss. Z. Humboldt-Univ. Berlin, Math. -Nat. Reihe*, 4 (1955) 1-7.
3. R. L. Kruse, On the circle group of a nilpotent ring, *this MONTHLY*, 77 (1970) 168-170.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

CROSSING NUMBER PROBLEMS

P. ERDÖS, Hungarian Academy of Science, and R. K. GUY, University of Calgary

A **graph**, $G(V, E)$, is a set V of **vertices** and a subset E of the unordered pairs of vertices, called **edges**. A **drawing** is a mapping of a graph into a surface. The vertices go into distinct points, **nodes**. An edge and its incident vertices map into a homeomorphic image of the closed interval $[0, 1]$ with the relevant nodes as end-

points and the interior, an **arc**, containing no node. A **good** drawing is one in which no two arcs incident with a common node have a common point; and no two arcs have more than one point in common. A common point of two arcs is a **crossing**. An **optimal** drawing in a given surface is one which exhibits the least possible number of crossings. Optimal drawings are good. This least number is the **crossing number** of the graph for the surface. We denote the crossing number of G for the plane (or sphere) by $v(G)$.

Almost all questions that one can ask about crossing numbers remain unsolved. For the **complete graph**, K_n , with n vertices and all $\binom{n}{2}$ possible edges, it has been conjectured [7] that

$$(1) \quad (?) \quad v(K_n) = \frac{1}{4} \left[\frac{1}{2}n \right] \left[\frac{1}{2}(n-1) \right] \left[\frac{1}{2}(n-2) \right] \left[\frac{1}{2}(n-3) \right],$$

where brackets denote greatest integer not greater than. For $n \leq 10$, this has been verified [10]:

n	2	3	4	5	6	7	8	9	10
$v(K_n)$	0	0	0	1	3	9	18	36	60

Blažek and Koman [1] and others [e.g., 7, 12] have given constructions which show that (1) is an upper bound. Kleitman's result [15, and see below] for the complete bipartite graph implies that for n sufficiently large,

$$(2) \quad v(K_n) \geq \frac{1}{80} n(n-1)(n-2)(n-3).$$

This is a little better than the lower bound given in [9]. It is easy to see that $v(K_n)/n^4$ is non-decreasing and so tends to a limit (between $\frac{3}{10}$ and $\frac{3}{8}$). A counting argument shows that if (1) is true for n odd, then it is also true for $n+1$. Eggleton and

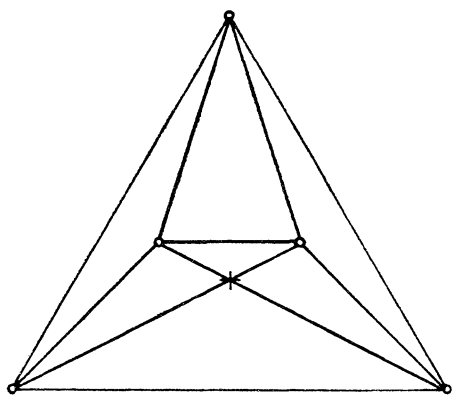


FIG. 1

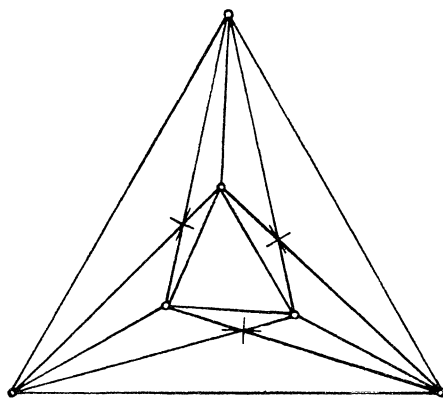


FIG. 2

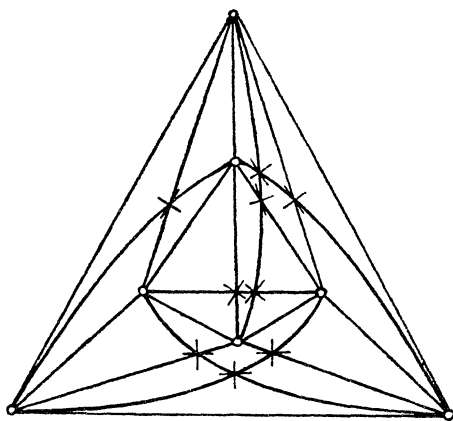


FIG. 3

Guy [3] have also shown that for n odd, $v(K_n)$ and $\binom{n}{4}$ have the same parity. Call two drawings **isomorphic** when there is a one-to-one correspondence between the nodes so that if any pair of arcs crosses, the corresponding pair also crosses. Optimal drawings of K_n for $n = 5, 6, 7, 8$ are shown in Figures 1, 2, 3, 4. For $n = 5, 6$ these are unique, but for $n = 7$ there are five which are non-isomorphic and for $n = 8$ there are three [10]. For $n = 9$ the number is about 200.

An attempt to put the theory of crossing numbers into algebraic form has been made by Tutte [20].

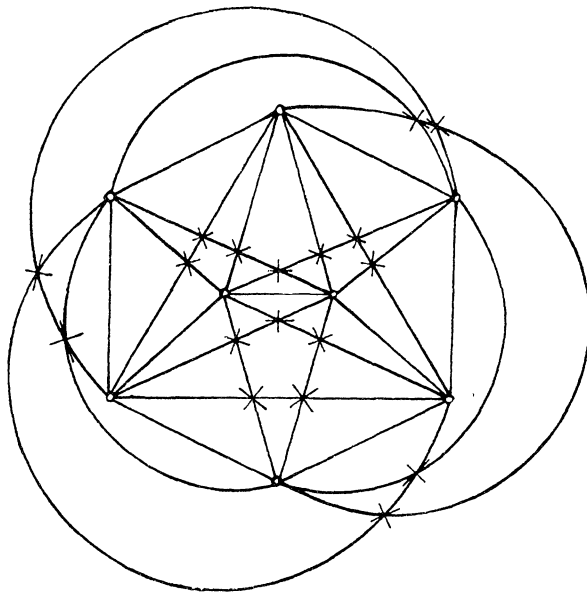


FIG. 4

If the arcs are restricted to be straight line-segments, we have the concept of **rectilinear crossing number**, $\bar{v}(G)$, of a graph G . It is clear that $\bar{v}(G) \geq v(G)$. A theorem of Fáry [6, 19] may be stated: if a graph can be embedded in the plane, then it can be so drawn using straight line segments. Hence $v(G) = 0$ implies $\bar{v}(G) = 0$. For $n \leq 7$ and $n = 9$, $v(K_n) = \bar{v}(K_n)$. (Figure 3 can be realized with straight line segments.) But Guy [10] has confirmed a conjecture of Harary and Hill [13] that $\bar{v}(K_8) = 19$, in contrast to $v(K_8) = 18$. It can also be shown that $\bar{v}(K_n) > v(K_n)$ for $n \geq 10$. It is conjectured that $\bar{v}(K_{10}) = 63$. Jensen [14] and independently Eggleton have shown that

$$(3) \quad \bar{v}(K_n) \leq [(7n^4 - 56n^3 + 128n^2 + 48n[(n-7)/3] + 108)/432]$$

and equality is conjectured. The fact that $\bar{v}(K_5) = 1$ gives an immediate proof of Esther Klein's result [5] that five points in the plane always include a convex quadrilateral. More generally, there is an exact correspondence between rectilinear crossings and convex quadrilaterals, so the problem of determining the rectilinear crossing number for the complete graph can be restated in the form: what is the least number of convex quadrilaterals determined by n points in the plane? More generally, one can ask for the least number of convex k -gons determined by n points in the plane, for $k > 4$. As before, the ratio of this number to $\binom{n}{k}$ tends to a positive limit as n tends to infinity with k fixed.

The crossing number problem for the **complete bipartite graph**, $K_{m,n}$, on $m+n$ vertices, whose mn edges are just those which join one of the m vertices to one of the n , first appeared as Turán's brick-factory problem. For some years it was thought that Zarankiewicz [22] and Urbaník [21] had solved this, but a hiatus in the proof was found independently by Ringel and Kainen [see 8] and the formula

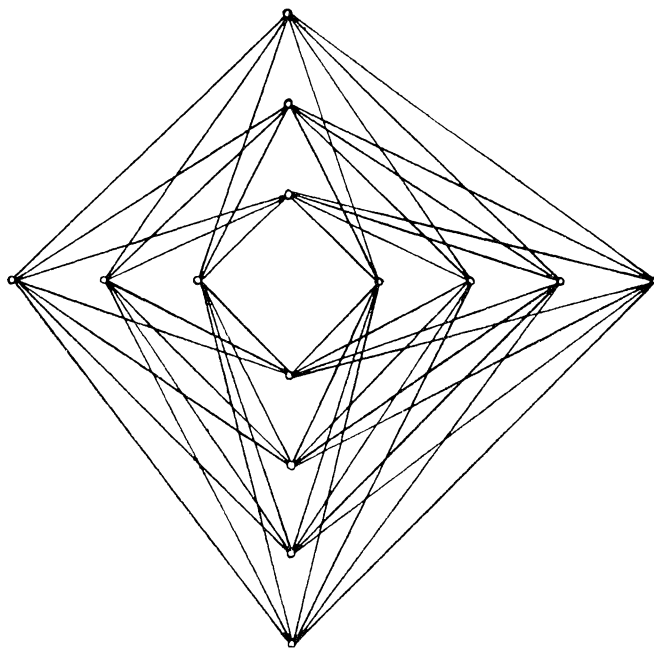
$$(4) \quad (?) \quad v(K_{m,n}) = \left\lfloor \frac{1}{2}m \right\rfloor \left\lfloor \frac{1}{2}(m-1) \right\rfloor \left\lfloor \frac{1}{2}n \right\rfloor \left\lfloor \frac{1}{2}(n-1) \right\rfloor$$

is still conjectural. It was established for $\min(m, n) = 3$ by Zarankiewicz and a counting argument again gives the result for each even number if it is known for the preceding odd one. The best result is due to Kleitman [15] who established (4) for $\min(m, n) \leq 6$. The corresponding rectilinear problem may have the same solution (4), since Zarankiewicz's construction uses only straight arcs (Figure 5).

For the 1-skeleton of the n -cube, Q_n , whose vertices, the 2^n binary n -tuples, are joined by an edge just if their vectors differ in exactly one component, Eggleton and Guy [4] announced that

$$(5) \quad (?) \quad v(Q_n) \leq \frac{5}{32}4^n - \left\lfloor \frac{n^2 + 1}{2} \right\rfloor 2^{n-2},$$

but a gap has been found in the description of the construction, so this must also remain a conjecture. We again conjecture equality in (5).



$$(?) \quad \bar{v}(K_{7,7}) = v(K_{7,7}) = 81.$$

FIG. 5

More generally, let $G(n, k)$ be a graph with n vertices and k edges. Denote by $g(n, k)$ the minimum of $v(G)$ taken over all graphs $G(n, k)$. Then we conjecture that

$$(6) \quad (?) \quad \frac{c_1 k^3}{n^2} < g(n, k) < \frac{c_2 k^3}{n^2};$$

in fact, that if $k/n \rightarrow \infty$, then $\lim g(n, k)/(k^3/n^2)$ exists. From Euler's theorem, $g(n, 3n-6) = 0$, $g(n, 3n-5) = 1$. The upper bound in (6) is trivial (with $c_2 = 1/8$), for, let l be the least integer with $ln > 2k$ and consider n/l copies of K_l . The lower bound would follow if we could prove that every drawing of a $G(n, k)$ contains an arc with at least $c_3 k^2/n^2$ crossings. In this connexion we can ask the following question: determine or estimate the smallest integer $f(r)$ so that every drawing of a graph $G(n, f(r))$ contains an arc with at least r crossings. Euler's theorem implies that $f(1) = 3n - 5$ and Eggleton and Guy [3] have shown that $f(2) = 4n - 8$ for $n = 6, 7$ and 9 , and $4n - 7$ for $n = 8$ or $n \geq 10$. This implies that

$$g(n, k) = k - 3n + 6 \text{ for } 3n - 6 \leq k \leq \min \left(4n - 8, \binom{n}{2} \right),$$

except that $g(7, 20) = 6$ and $g(9, 28) = 8$. But $f(3)$ has not yet been determined.

Another related question is: which graphs $G(n, k)$ have maximal $v(G)$ and what

is this maximum? We conjecture that the following graph has maximal $v(G)$: take l so that

$$\binom{l}{2} \leq k < \binom{l+1}{2},$$

and the graph consists of K_l with a vertex joined to $k - \binom{l}{2}$ of its vertices (and $n - l - 1$ isolated points).

These more general problems can also be posed in the rectilinear case. We can also ask analogous questions for surfaces of higher genus; some results have been obtained for the torus [11, 12], and for the projective plane and Klein bottle [16].

We are indebted to R. B. Eggleton for helpful discussions and suggestions, and permission to reproduce his results.

References

1. J. Blažek and M. Koman, A minimal problem concerning complete plane graphs, in M. Fiedler (ed.), *Theory of Graphs and its Applications*, Proc. Symp. Smolenice, 1963; Prague, 1964, 113–117; MR 30(1965) #4249.
2. ———, and ———, On an extremal problem concerning graphs, *Comm. Math. Univ. Carolinae*, 8(1967) 49–52; MR 35(1968) #1506.
3. R. B. Eggleton, Ph.D. thesis, Univ. of Calgary, 1973.
4. R. B. Eggleton and R. K. Guy, The crossing number of the n -cube, *Amer. Math. Soc. Notices*, 17(1970) 757.
5. P. Erdős and G. Szekeres, A combinatorial problem in geometry, *Compositio Math.*, 2 (1935) 463–470.
6. I. Fáry, On straight line representation of planar graphs, *Acta Sci. Math. (Szeged)*, 11 (1948) 229–233; MR 10(1949) 136.
7. R. K. Guy, A combinatorial problem, *Nabla (Bull. Malayan Math. Soc.)* 7 (1960) 68–72.
8. ———, The decline and fall of Zarankiewicz's theorem, in F. Harary (ed.), *Proof Techniques in Graph Theory*, Academic Press, N. Y., 1969, 63–69.
9. ———, Sequences associated with a problem of Turán and other problems, *Proc. Balatonfüred Combinatorics Conf.*, 1969. Bolyai János Matematikai Társulat, Budapest, 1970, 553–569.
10. ———, Latest results on crossing numbers, in *Recent Trends in Graph Theory*, Springer, N.Y., 1971, 143–156.
11. R. K. Guy and T. A. Jenkyns, The toroidal crossing number of $K_{m,n}$, *J. Combinatorial Theory*, 6 (1969) 235–250; MR 38(1969) #5660.
12. R. K. Guy, T. A. Jenkyns and J. Schaer., The toroidal crossing number of the complete graph, *J. Combinatorial Theory*, 4(1968) 376–390, MR 36(1968) #3682.
13. F. Harary and A. Hill, On the number of crossings in a complete graph, *Proc. Edinburgh Math. Soc.* (2), 13 (1962–3) 333–338; MR 29 (1965) #602.
14. H. F. Jensen, An upper bound for the rectilinear crossing number of the complete graph, *J. Combinatorial Theory*, 10B (1971) 212–216.
15. D. J. Kleitman, The crossing number of $K_{5,n}$, *J. Combinatorial Theory*, 9 (1970) 315–323.
16. M. Koman, On the crossing numbers of graphs, *Acta, Univ. Carolinae Math. Phys.*, 10 (1969) 9–46.
17. K. Kuratowski, Sur le problème des courbes gauches en topologie, *Fund. Math.*, 15 (1930) 271–283.

18. J. W. Moon, On the distribution of crossings in random complete graphs, *J. Soc. Indust. App. Math.*, 13 (1965) 506–510; MR 31 (1966) #3357.

19. W. T. Tutte, How to draw a graph, *Proc. London Math. Soc.* (3), 13 (1963) 743–767; MR 28 (1964) #1610.

20. ———, Towards a theory of crossing numbers, *J. Combinatorial Theory*, 8 (1970) 45–53.

21. K. Urbanik, Solution du problème posé par P. Turán, *Colloq. Math.*, 3 (1955) 200–201.

22. K. Zarankiewicz, On a problem of P. Turán concerning graphs, *Fund. Math.*, 41 (1954) 137–145; MR 16 (1955) 156.

CLASSROOM NOTES

EDITED BY ROBERT GILMER

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306. Notes are usually limited to three printed pages.

A PROOF OF UNIQUENESS OF FACTORIZATION IN THE GAUSSIAN INTEGERS

M. F. RUCHTE AND R. W. RYDEN, Humboldt State College

Let $K(i)$ denote the Gaussian Integers, $K(i) = \{a + bi \mid a, b \text{ are rational integers}\}$. It is well known that $K(i)$ has the unique factorization property. Normally, one shows that $K(i)$ is a Euclidean domain and then uses the fact that every Euclidean domain is a unique factorization domain. We give a direct proof that factorization is unique in $K(i)$ which parallels the proof for the rational integers as given in Niven and Zuckerman (p. 15). We would like to express our appreciation to Professor Ivan Niven for having raised for us the question of the existence of this type of proof.

LEMMA 1. *If z and w are two non-zero complex numbers such that $|w| \leq |z|$ and $|\arg z - \arg w| < \pi/3$, then $|z - w| < |z|$.*

Proof. The triangle formed by the points 0, z , w in the complex plane has an angle less than $\pi/3$ at the origin, so the side opposite, which is of length $|z - w|$, cannot be the longest side. Further, since $|w| \leq |z|$, we conclude that $|z - w| < |z|$.

If z is a complex number the associates of z are the numbers z , $-z$, iz , $-iz$.

LEMMA 2. *If z and w are complex numbers then there exists an associate w' of w such that $|\arg z - \arg w'| < \pi/3$.*

Proof. The associates of w are at right angles to one another; therefore, there must be one of them in any given sector of angle $2\pi/3$.

If $\alpha \in K(i)$, $\alpha = a + bi$, denote by $N(\alpha)$, the norm of α , the non-negative rational integer $a^2 + b^2$. Note that (1) $N(\alpha\beta) = N(\alpha) \cdot N(\beta)$, (2) if ε is a unit ($\varepsilon = 1, -1, i, \text{ or } -i$) then $N(\varepsilon) = 1$, and (3) $N(\alpha) = |\alpha|^2$.

Two factorizations $\alpha \cdot \beta, \mu \cdot \delta$ are the same up to associates if α and β are respectively associates of the μ and δ in either order and similarly for factorizations into more than two factors.

THEOREM. *Factorization in $K(i)$ is unique up to associates.*

Proof. Suppose that $K(i)$ does not have unique factorization and that α is a smallest number in norm which admits two representations,

$$\alpha = \sigma_1 \cdots \sigma_r = \rho_1 \cdots \rho_s \quad (r, s > 1),$$

where no σ_i is an associate of any ρ_j . Without loss of generality we may assume that $|\sigma_1| \geq |\rho_1|$ and, since the factorization can be changed by taking associates of the involved primes, that $|\arg \sigma_1 - \arg \rho_1| < \pi/3$ by Lemma 2.

Let $\beta = (\sigma_1 - \rho_1)\sigma_2\sigma_3 \cdots \sigma_r$,

$$N(\beta) = N(\sigma_1 - \rho_1)N(\sigma_2) \cdots N(\sigma_r) < N(\sigma_1)N(\sigma_2) \cdots N(\sigma_r) = N(\alpha).$$

But

$$\begin{aligned} \beta &= \sigma_1\sigma_2 \cdots \sigma_r - \rho_1\sigma_2 \cdots \sigma_r \\ &= \rho_1 \cdots \rho_s - \rho_1\sigma_2 \cdots \sigma_r \\ &= \rho_1(\rho_2 \cdots \rho_s - \sigma_2 \cdots \sigma_r). \end{aligned}$$

So that β admits two representations, one having ρ_1 as a factor and one having no associate of ρ_1 as a factor; a contradiction of the minimality of $N(\alpha)$.

Notice that geometric arguments can be applied to all complex quadratic fields. Lemma 2, however, relies on the existence of four associates at right angles. For the case $K(\sqrt{-3})$ we also have enough units (see Niven and Zuckerman, p. 210) to yield the conclusion of Lemma 2 so that unique factorization for $K(\sqrt{-3})$ follows in the same fashion. However, all other $K(\sqrt{m})$ where m is negative, have only two units, ± 1 , so that this particular proof cannot be extended to the case of other negative m .

Reference

1. I. Niven and H. Zuckerman, *An Introduction to the Theory of Numbers*, 2nd ed., Wiley, New York, 1966.

SOME HALF-PLANE DIRICHLET PROBLEMS: A BARE HANDS APPROACH

F. J. FLANIGAN, University of California, San Diego

1. We suggest that students in a basic complex variables course might benefit from seeing, early in the course, that analytic functions give ready-made solutions to boundary-value problems, and we offer a class of attractive Dirichlet problems for the upper half plane to illustrate this fact. These problems require only (a) the

theorem that the real and imaginary parts of a complex analytic function are harmonic, and (b) a simple version of the partial fractions decomposition of a rational function, as seen in calculus. This project could be arranged in exercise form, soon after the presentation of the Cauchy-Riemann equations. The student will be delighted to discover that he does not have to know here "how to solve partial differential equations," for this is taken care of by (a) above.

2. The problem: Given the "boundary values" $B(x)$, a real-valued rational function defined everywhere on the x -axis, to find a real-valued function $H(x, y)$ continuous on the closed upper half plane $y \geq 0$ and harmonic on the open upper half plane $y > 0$ such that $H(x, 0) = B(x)$. Note therefore that $B(x) = P(x)/Q(x)$ where P and Q are real polynomials with no common roots, and $Q(x)$ has no real roots.

Our method is always to find the "correct" complex analytic function and use its real part as the solution $H(x, y)$.

3. First case: $B(x)$ is a polynomial in x . Here $H(x, y) = \operatorname{Re} B(x + iy)$ is a solution. (We do not discuss uniqueness.)

However, if $B(x) = 1/(x^2 + 1)$, then $\operatorname{Re} B(x + iy)$ is *not* a solution, because of the pole at $z = i$. The solution in this case is, in fact,

$$H(x, y) = 2 \operatorname{Re} \left[\frac{i}{2} \cdot \frac{1}{z + i} \right] = \frac{y + 1}{x^2 + (y + 1)^2}.$$

This comes from the following considerations.

4. We consider only $B(x) = P(x)/Q(x)$ where the roots of $Q(x)$ are non-real and simple.

LEMMA. Given $B(x)$ as above, there is a decomposition

$$B(x) = B^+(x) + B^-(x),$$

where B^+, B^- are rational functions of x with complex coefficients such that

- (i) the values $B^+(x), B^-(x)$ are conjugate complex numbers,
- (ii) the complexified functions $B^+(z), B^-(z)$ have their poles concentrated in the lower half plane and upper half plane, respectively.

Proof (outline). (a) As remarked above, we assume that $Q(x)$ has roots z_1, \dots, z_n (open upper half plane) and $\bar{z}_1, \dots, \bar{z}_n$, all distinct.

(b) We write down a formal partial fraction decomposition

$$1/Q(x) = \sum_j [a_j/(x - z_j)] + \sum_j [b_j/(x - \bar{z}_j)]$$

and argue that equality holds provided

$$a_j = [(z - z_j)/Q(z)]_{z=z_j}, \quad b_j = [(z - \bar{z}_j)/Q(z)]_{z=\bar{z}_j}.$$

One uses here the hypothesis that the roots of $Q(z)$ are distinct.

(c) One next observes $b_j = \bar{a}_j$.

(d) But since $B(x) = P(x)/Q(x)$, we have

$$B(x) = P(x) \sum_j [a_j/(x - z_j)] + P(x) \sum_j [\bar{a}_j/(x - \bar{z}_j)].$$

We name the two terms separated by the $+$ sign here $B^-(x)$, $B^+(x)$ and verify immediately that they satisfy (i) and (ii). Done.

Just as in Section 3, we may now write down a solution to the Dirichlet problem, namely

$$H(x, y) = 2 \operatorname{Re} B^+(x + iy).$$

The author acknowledges support from the National Science Foundation through NSF GP-23104.

SINGLE LAYER POTENTIALS AND THE CAUCHY-KOWALEWSKI THEOREM

P. A. NICKEL, North Carolina State University

Single and double layer potentials have played a very important role in mathematical physics [1] for a long time and an important problem related to these is the determination of the jumps in the potentials themselves, as well as the jumps in the associated normal and tangential forces. In standard treatments of these phenomena such as [2] and [3], arguments of an advanced calculus sort are used, and these become extremely delicate.

In his book on partial differential equations [4], Garabedian makes a convincing case for the use of the Cauchy-Kowalewski Theorem in discussing jump phenomena in these problems. The reason for this is quite simple; namely, when the path of integration is deformed and an equivalent distribution defined, the integrands become regular. As a result, the arguments of advanced calculus are routine, at least when the boundary curve $\partial\mathcal{D}$ and distributions are analytic.

1. Purpose of this note. In the determination of the jump in the normal force for a single layer potential in [4], an argument is made concerning the symmetry of the contributions from two sections of a certain spherical surface in the limit as the radius of the sections goes to 0. Even though this phrase is reasonable enough, it would seem desirable to establish the jump relations from the inside and outside of \mathcal{D} independently of one another, as was done in the discussion of continuity of a single layer potential, as well as in the discussion of continuity of the normal force for a double layer. The purpose then, of this note is the determination of the normal force for a single layer μ at points of $\partial\mathcal{D}$ inside \mathcal{D} without recourse to the argument mentioned above concerning symmetry in the limit. The same can then

be accomplished for points on $\partial\mathcal{D}$ outside \mathcal{D} , and the jump is simply the difference of these forces.

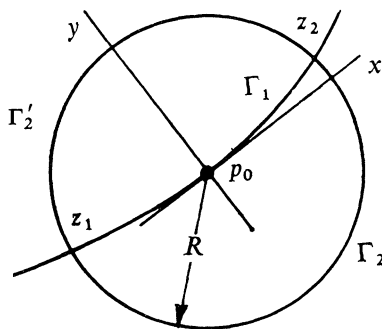
2. Notation. The presentation here is made in two dimensions, at a point p_0 on the analytic boundary of the simply connected region \mathcal{D} . The normal force at $p_0 \in \partial\mathcal{D}$ inside of \mathcal{D} is

$$\lim_{p \rightarrow p_0} (\partial V^- / \partial n) = (\partial V^- / \partial n)_{p_0}.$$

In particular, our task is then to develop the relation (9.46') of [4]

$$(1) \quad \left. \frac{\partial V^-}{\partial n} \right|_{p_0} = -\pi\mu(p_0) + \int_{\partial\mathcal{D}} \mu(s) \left. \frac{\partial}{\partial n} \log r \right|_{p_0} ds,$$

where $(\partial/\partial n)$ refers to differentiation in x and y in the direction of the outward normal and $r = \sqrt{(x - \xi_s)^2 + (y - \eta_s)^2}$, the distance from (x, y) to the point (ξ_s, η_s) on $\partial\mathcal{D}$, all in terms of arc length s .



For convenience, we select p_0 as the origin and take the x -axis in the direction of the tangent to $\partial\mathcal{D}$ at p_0 . Hence there is a disc $\Delta_{R_1}(p_0)$ of radius R_1 and center at p_0 inside of which $\partial\mathcal{D}$ is described by $y = f(x)$, with $f(0) = f'(0) = 0$. Further, with an application of the Cauchy-Kowalewski Theorem as in [4, p. 336], there is another disc, say $\Delta_{R_2}(p_0)$ inside of which there is a harmonic function $u(z)$ such that on $\Delta_{R_2}(p_0) \cap \partial\mathcal{D}$

$$(2) \quad u = 0, \quad \frac{\partial u}{\partial n} = \mu.$$

Hence we can employ the conditions (2) as well as the description $y = f(x)$ of $\partial\mathcal{D}$ inside of $\Delta_R(p_0)$ where $R = \min(R_1, R_2)$. Evidently the boundary $\partial\Delta_R(p_0)$ meets $\partial\mathcal{D}$ in exactly two points, denoted z_1 and z_2 , with arguments taken as $\pi - \theta_1$ and $2\pi + \theta_2$. (The figure illustrates θ_1 and θ_2 as positive, but this condition is not vital to the argument.)

Applying Green's identity to the region bounded by Γ_1 and Γ_2 , along with (2),

we convert the integral for V from an integral along $\Gamma_1 + \partial\mathcal{D} - \Gamma_1$ to the integral

$$(3) \quad V = \int_{\partial\mathcal{D}-\Gamma_1} \mu \log r ds + \int_{\Gamma_2} \frac{\partial u}{\partial v} \log r \Big|_{(\xi_s, \eta_s)} ds - \int_{\Gamma_2} u \frac{\partial}{\partial v} \log r \Big|_{(\xi_s, \eta_s)} ds.$$

Here, $(\partial/\partial v)$ again represents differentiation in the direction of the outward normal, but in the dummy variables (ξ, η) , and $\Gamma_2 = \{z; |z| = R \text{ and } \arg z_1 \leq \arg z \leq \arg z_2\}$.

3. The normal force. Since the path of integration no longer passes through p_0 , the normal differentiations are performed before integration rather than after, and the task of establishing (1) is that of determining I_y and J_y , the partial derivatives of the second and third integrals of (3) in the coordinate system with origin at p_0 . If on Γ_2 , ξ_s and η_s are written in polar coordinates (ρ, θ) , we find in terms of θ_1 and θ_2 ,

$$\begin{aligned} \frac{\partial I}{\partial y} \Big|_{(0,0)} &= \int_{\pi-\theta_1}^{2\pi+\theta_2} \frac{\partial u}{\partial \rho} \frac{\partial}{\partial y} \log \sqrt{(x-R\cos\theta)^2 + (y-R\sin\theta)^2} \Big|_{(0,0)} R d\theta \\ (4) \quad &= - \int_{\pi-\theta_1}^{2\pi+\theta_2} \frac{\partial u}{\partial \rho} \Big|_{\rho=R} \sin\theta d\theta \\ &= - \int_{\pi-\theta_1}^{2\pi+\theta_2} (u_x(0,0)\cos\theta + u_y(0,0)\sin\theta) \sin\theta d\theta + I_R, \end{aligned}$$

where I_R is an integral which goes to 0 with R . But θ_2 and θ_1 go to 0 with R as well, for $\tan\theta_2 = y_2/x_2 = \frac{1}{2}x_2 f''(\bar{x})$, with $0 < \bar{x} < x_2$. But f'' is bounded and it follows that $\tan\theta_2 \rightarrow 0$ as $x_2 \rightarrow 0$. But as $R \rightarrow 0$, certainly $x_2 \rightarrow 0$. Hence θ_2 , itself selected in the range $(-\frac{1}{2}\pi, \frac{1}{2}\pi)$, must go to 0 as well. Recalling that $u_x(0,0) = 0$ by virtue of (2), and letting $R \rightarrow 0$, we find $I_y(0,0) = -(\pi/2)u_y(0,0) = (\pi/2)\mu(s_0)$.

The last term of (3) is handled in the same way:

$$\begin{aligned} (5) \quad \frac{\partial J}{\partial y} \Big|_{(0,0)} &= \frac{\partial}{\partial y} \int_{\Gamma_2} u \frac{\partial}{\partial v} \log \sqrt{(x-\xi_s)^2 + (y-\eta_s)^2} ds \Big|_{(0,0)} \\ &= \int_{\pi-\theta_1}^{2\pi+\theta_2} u \sin\theta \frac{d\theta}{R} = \int_{\pi-\theta_1}^{2\pi+\theta_2} (u_x(0,0)\cos\theta + u_y(0,0)\sin\theta) \sin\theta d\theta + O(R) \end{aligned}$$

and, in the limit $J_y(0,0)$ is just $-(\pi/2)\mu(s_0)$. The equation (1) now follows from (3) and the relation $(\partial/\partial n)_{p_0} = -(\partial/\partial y)_{(0,0)}$ and our purpose is now achieved.

As a further observation, we see that the force $(\partial V^+/\partial n)_{p_0}$, represented by replacing the path Γ_1 by Γ'_2 , is in the limit as $R \rightarrow 0$ just the negative of the corresponding contribution to $(\partial V^-/\partial n)_{p_0}$ from Γ_2 . For the only difference in the analysis is that the normal derivatives on Γ'_2 will be $-(\partial/\partial \rho)$, and the non-symmetric contributions are $O(R)$. Hence, $(\partial V^+/\partial n)_{p_0}$ and $(\partial V^-/\partial n)_{p_0}$, except for a difference in sign, can differ only by $O(R)$; that is, we are back to the statement of [4, p. 339] that "these contributions can differ only in sign in the limit as $R \rightarrow 0$."

Acknowledgement.

The author expresses his thanks to the referee for his suggestions and, in particular, for his questioning one very important sign.

References

1. Julius A. Stratton, *Electromagnetic Theory*, McGraw-Hill, New York, 1941.
2. O. D. Kellogg, *Foundations of Potential Theory*, Ungar, New York, 1946.
3. I. G. Petrovskii, *Partial Differential Equations*, Saunders, Philadelphia, 1967.
4. P. R. Garabedian, *Partial Differential Equations*, Wiley, New York, 1964.

A GLOBAL CHARACTERIZATION OF UNIFORM CONTINUITY

RICHARD CLEVELAND, Sacramento State College

Let f be a function on a metric space (X, d) into a metric space (Y, D) . It is well known that f is uniformly continuous on X if and only if for any two sequences $\{x_n\}$ and $\{y_n\}$ in X ,

$$\lim_n d(x_n, y_n) = 0 \text{ implies } \lim_n D(f(x_n), f(y_n)) = 0,$$

[e.g., [1], p. 168]. From this one easily obtains the following global property of uniform continuity.

THEOREM 1. *If f is uniformly continuous on X and A and B are non-empty subsets of X with $d(A, B) = 0$, then $D(f[A], f[B]) = 0$.*

The converse of this theorem is also true; the property of preserving zero distance between sets characterizes uniform continuity. This fact was first announced by Yu. M. Smirnov in [3]. It later appeared in a paper by H. Kenyon [2]. Both of these papers give the result in the setting of general uniform spaces. It is hoped that this note will make it more accessible to students and teachers.

THEOREM 2. *If for every pair of non-empty subsets A and B of X*

$$(1) \quad d(A, B) = 0 \text{ implies } D(f[A], f[B]) = 0,$$

then f is uniformly continuous on X .

Proof. Suppose (1) holds, but f is not uniformly continuous on X . Then choose $\varepsilon > 0$ so that

(2) for every $\delta > 0$ there are x and y in X such that $d(x, y) < \delta$ and $D(f(x), f(y)) \geq 3\varepsilon$.

We keep ε fixed for the rest of the proof. For any $z \in X$, let

$$S(z) = \{s \in X : D(f(s), f(z)) < \varepsilon\}$$

and

$$T(z) = \{t \in X : D(f(t), f(z)) \geq 2\varepsilon\}.$$

Notice that for any z , if $T(z) \neq \emptyset$,

$$D(f[S(z)], f[T(z)]) \geq \varepsilon,$$

so that by (1),

$$d(S(z), T(z)) > 0.$$

We choose inductively a sequence $\{\delta_n\}$ of positive numbers and sequences $\{x_n\}$ and $\{y_n\}$ in X as follows: take $\delta_1 = 1$ and choose x_1 and y_1 by (2) so that

$$d(x_1, y_1) < \delta_1 \text{ and } D(f(x_1), f(y_1)) \geq 3\varepsilon.$$

Suppose δ_n , x_n , and y_n have been chosen. Then take δ_{n+1} so that

$$\begin{aligned} 0 &< \delta_{n+1} \\ &< \min\{\tfrac{1}{2}\delta_n, d(S(x_n), T(x_n)), d(S(y_n), T(y_n))\} \end{aligned}$$

and choose x_{n+1} and y_{n+1} in X so that

$$d(x_{n+1}, y_{n+1}) < \delta_{n+1}$$

and

$$D(f(x_{n+1}), f(y_{n+1})) \geq 3\varepsilon.$$

Now let $A = \{x_1, x_2, \dots\}$ and $B = \{y_1, y_2, \dots\}$. By construction, $\lim_n \delta_n = 0$, so $d(A, B) = 0$. The proof will be complete when we show that

$$D(f[A], f[B]) > 0.$$

To this end, suppose m and n are integers and $m < n$. Then $d(x_n, y_n) < \delta_n$ and $D(f(x_n), f(y_n)) \geq 3\varepsilon$. Suppose $D(f(x_n), f(y_m)) < \varepsilon$. This means that $x_n \in S(y_m)$. But since $d(S(y_m), T(y_m)) > \delta_n > d(x_n, y_n)$, it follows that $y_n \notin T(y_m)$, or

$$D(f(y_n), f(y_m)) < 2\varepsilon.$$

But then we obtain the contradiction

$$D(f(x_n), f(y_n)) \leq D(f(x_n), f(y_m)) + D(f(y_n), f(y_m)) < 3\varepsilon.$$

Therefore,

$$D(f(x_n), f(y_m)) \geq \varepsilon.$$

By a similar argument we obtain the same inequality for $m > n$, and since we already have it for $m = n$, we have shown that

$$D(f[A], f[B]) \geq \varepsilon,$$

and the proof is complete.

It is clear from the proof that the hypothesis of this theorem can be weakened to the assumption that (1) holds for every pair of countable subsets of X .

To see how this theorem can be used as a test for uniform continuity, consider the following applications:

THEOREM 3. *If a sequence of uniformly continuous functions converges uniformly on X , then the limit function is uniformly continuous.*

Proof. Suppose $f_n \rightarrow g$ uniformly on X , where each f_n is uniformly continuous on X . Let A and B be non-empty subsets of X with $d(A, B) = 0$. Let $\varepsilon > 0$ and choose n so large that

$$D(f_n(t), g(t)) < \varepsilon/3,$$

for all $t \in X$. Then choose $x \in A$ and $y \in B$ so that

$$D(f_n(x), f_n(y)) < \varepsilon/3.$$

Then it follows that

$$D(g(x), g(y)) \leq D(g(x), f_n(x)) + D(f_n(x), f_n(y)) + D(f_n(y), g(y)) < \varepsilon,$$

and we conclude that $D(g[A], g[B]) = 0$.

THEOREM 4. *If X is compact and f is continuous on X , then f is uniformly continuous on X .*

Proof. Suppose A and B are non-empty subsets of X such that $d(A, B) = 0$. Because X is compact, this implies that $A^- \cap B^- \neq \emptyset$. Also, because f is continuous, $f[S^-] \subset f[S]^-$ for any $S \subset X$. Combining these two facts, we get

$$\emptyset \neq f[A^-] \cap f[B^-] \subset f[A]^- \cap f[B]^- ,$$

so that $D(f[A], f[B]) = 0$.

References

1. R. G. Bartle, *Elements of Real Analysis*, New York, 1964.
2. H. Kenyon, Two theorems on relations, *Trans. Amer. Math. Soc.*, 107 (3) (1963) 1-14.
3. Yu. M. Smirnov, On proximity spaces, *Mat. Sb.*, (N. S.) 31 (73) (1952) 543-574.

A LETTER BY PROFESSOR POLYA

The following letter was written recently by Professor Polya to a Department Chairman. It is reproduced here with Professor Polya's permission, but with the omission of any other identifying names.

"Dear Colleague:

"As you may know, I am especially concerned with problem-solving; I wrote books about it and I stressed it in my teaching, especially in teaching high school mathematics teachers. That the role of problem-solving in mathematics is not understood by non-mathematicians and is not duly appreciated by outsiders, is not surprising and we need not worry about it. But I heard lately that such lack of understanding and appreciation led to denying the promotion to a member of your Department. I feel that there is a serious matter of principle involved, and I wish to write you about it.

"Problems play an essential role both in the progress and in the teaching of science. I cannot develop properly this topic in this letter — it would need two volumes, one on history and methodology, another on pedagogy. Yet let us come nearer to the particular case we are concerned with.

"Problems play an important role on all levels of mathematical instruction. It is by solving problems that the students learn to understand, to apply and to appreciate the material presented in the course, and the instructor judges the performance of the students on the basis of their problem-solving. More advanced students may do some other kind of work (e. g., reports in a seminar) but problems are the backbone of undergraduate instruction.

"Demands on the teacher are different on different levels.

"A faculty member who teaches mainly graduate students must prepare them for research and so he has the duty to keep contact with contemporary research and cannot let his own research get rusty. Does he do his duty? How can we judge it? Most directly on the basis of his publications. It is well known that the 'principle' of 'publish or perish' was unwisely applied in several cases, yet some rule in this direction is necessary to judge the instructors of advanced students.

"A faculty member who teaches mainly undergraduate students should have, of course, a good mathematical background and he should not let it get rusty. Yet to extend to his case the 'principle' of 'publish or perish' is unwise and unjust. Under stress — and just for prestige, without real love or interest, the faculty member finally produces a paper that is printed and immediately submerged, unread and unnoticed, in the ocean of the present overproduction — is not such an effort misguided? Another way of not getting rusty is to pose and solve problems — and it is, in my opinion, in many cases a better way: Problem-

solving is a perfectly acceptable and respectable professional activity for a mathematician and can favorably influence his teaching. The problem section of the *American Mathematical Monthly*, e.g., contains some quite difficult problems, has very good editorial staff, and appeals to a good number of problem-solving mathematicians, some of whom are quite enthusiastic.

"If it is true what I heard that your colleague's promotion was refused, because he 'only' solved problems and did not publish, such a decision is unwise and unjust.

"Sincerely yours

(Signed by)

"GEORGE POLYA

Professor Emeritus, Stanford University"

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

ASSOCIATE EDITORS: JOSHUA BARLAZ, ERIC S. LANGFORD. COLLABORATING EDITORS: LEONARD CARLITZ, GULBANK D. CHAKERIAN, HASKELL COHEN, S. ASHBY FOOTE, ISRAEL N. HERSTEIN, MURRAY S. KLAMKIN, DANIEL J. KLEITMAN, ROGER C. LYNDON, MARVIN MARCUS, CHRISTOPH NEUGEBAUER, ALBERT WILANSKY, AND UNIVERSITY OF MAINE PROBLEMS GROUP: GEORGE S. CUNNINGHAM, CLAYTON W. DODGE, HOWARD W. EVES, WILLIAM R. GEIGER, GARY HAGGARD, PHILIP M. LOCKE, JOHN C. MAIRHUBER, CURTIS S. MORSE, GRATTAN P. MURPHY, EDWARD S. NORTHAM AND WILLIAM L. SOULE, JR.

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before April 30, 1973. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.

E 2391. Proposed by V. R. R. Uppuluri, Oak Ridge National Laboratory

It is well known that three chords can divide a circular disk into at most seven pieces. Can these seven pieces all have the same area?

E 2392. *Proposed by David Singmaster, Polytechnic of the South Branch, London, England*

On the $n \times n$ chessboard, for $n \geq 4$, define the *knight's distance* $D(A, B)$ between the squares A and B to be the minimum number of knight's moves required to go from A to B . Define the *knight's diameter* $M(n)$ of the $n \times n$ board to be the maximum knight's distance between any two squares on the board.

1. Is $M(n)$ monotonic?
2. Does $M(n)$ always equal the knight's distance between opposite corners of the board?
3. Prove or disprove: For $n \geq 5$, $M(n) = \lceil 2n/3 \rceil$.
4. Determine the knight's distance $D(0, P)$ from the origin to an arbitrary square $P = (a, b)$ on the infinite chessboard.

E 2393. *Proposed by M. S. Klamkin, Ford Motor Company*

Parallel lines are drawn through the vertices A_0, A_1, \dots, A_n of a given simplex of volume V , terminating in the opposite faces (extended if necessary) in the points B_0, B_1, \dots, B_n , respectively.

(1) Show that the volume of the simplex determined by B_0, B_1, \dots, B_n is nV .

(2) Show that the volume of the simplex determined by the vertices $A_0, A_1, \dots, A_r, B_{r+1}, B_{r+2}, \dots, B_n$ is given by $V_r' = \left| \frac{n-r-1}{n} \right| V$.

E 2394. *Proposed by S. L. Greitzer, Rutgers University, and M. S. Klamkin, Ford Motor Company*

A line is drawn through the centroid G of a simplex A_0, A_1, \dots, A_n intersecting the faces (extended if necessary) in the points B_0, B_1, \dots, B_n , respectively. Show that

$$\sum_{i=0}^n \frac{1}{GB_i} = 0,$$

where GB_i denotes the directed distance from G to B_i . Show also that the above property characterizes the point G as the centroid; i.e., if the above sum vanishes for all arbitrary lines, then G is the centroid.

This generalizes known results for triangles and tetrahedrons.

E 2395. *Proposed by H. W. Gould, West Virginia University*

Let n be a nonnegative integer. For $p = 1, 2, \dots$ define

$$A_p(n) = \sum_{0 \leq k \leq n/2} (-1)^k \left\{ \binom{n}{k} - \binom{n}{k-1} \right\}^p,$$

where we make the usual conventions regarding binomial coefficients. Prove that, whenever n is odd, $A_2(n) = nA_1(n)$.

E 2396. *Proposed by Erwin Just, Bronx Community College*

(A) Prove that $2^p + 3^p$ is not a perfect power if p is prime.

(B) Find all natural numbers m and n such that $2^m + 3^n$ is a perfect square.

SOLUTIONS OF ELEMENTARY PROBLEMS

Enumeration via the Chinese Remainder Theorem

E 2330 [1971, 1138]. *Proposed by Richard Stanley, Massachusetts Institute of Technology*

Let f be a function from the positive integers to the integers satisfying $f(m+n) \equiv f(n) \pmod{m}$ for all $m, n \geq 1$ (e.g., a polynomial with integer coefficients). Let $g(n)$ be the number of values (including repetitions) of $f(1), f(2), \dots, f(n)$ divisible by n , and let $h(n)$ be the number of these values relatively prime to n . Show that g and h are multiplicative functions of n related by

$$h(n) = \sum_{d|n} \mu(d)g(d)(n/d) = n \prod_{p|n} \left(1 - \frac{g(p)}{p}\right).$$

Solution by Stephen Spindler, University of Chicago. Given $(m, n) = 1$ and $1 \leq a \leq m, 1 \leq b \leq n$, it follows from the Chinese Remainder Theorem and the properties of f that $m \mid f(a)$ and $n \mid f(b)$ if and only if $mn \mid f(x)$ where $x = x(a, b)$ is that unique integer such that $x \equiv a \pmod{m}, x \equiv b \pmod{n}$, and $1 \leq x \leq mn$. Thus g is multiplicative. For $d \mid n$, the number of values of $f(1), \dots, f(n)$ divisible by d is just $(n/d)g(d)$; by a straightforward inclusion-exclusion count,

$$h(n) = n - \sum (n/p)g(p) + \sum (n/pp')g(pp') - \dots,$$

the first sum being over all primes p such that $p \mid n$, the second being over all pairs of distinct primes p, p' such that $pp' \mid n$, etc. Thus

$$h(n) = n \prod_{p|n} \left(1 - \frac{g(p)}{p}\right),$$

as desired.

Also solved by Arnold Adelberg, S. J. Benkoski, Suzette M. Cormier, Neal Felsing, J. L. Goodling, M. G. Greening (Australia), Emil Grosswald, H. S. Hahn, J. L. Hunsucker & Jack Nebb, Wells Johnson, David Kelly, J. F. Marcotorchino (France), L. E. Mattics, Kenneth Rosen, Edward Rosenthal, Temple University Problem Solving Group, and the proposer.

Editor's Comment: Benkoski calls attention to Harlan Stevens, *Generalizations of the Euler ϕ -function*, Duke Math. J., 38 (1971), 181–186, and shows how minor modifications of the results of this paper can be used to solve the present problem. The proposer notes that the condition $f(m+n) \equiv f(n) \pmod{m}$ can be replaced by the more general condition that m divides $f(n)$ if and only if m divides $f(m+n)$.

Divers Diverse Diophantine Determinations

E 2332 [1972, 87]. *Proposed by R. S. Luthar, University of Wisconsin at Janesville*

Find all solutions in positive integers:

$$y^3 + 4y = z^2.$$

I. *Solution by G. B. Robinson, State University of New York at New Paltz.* Let $y = nk^2$ with n square-free. Since $k^2 \mid z^2$, it follows that $z = mk$, reducing the equation to $n(y^2 + 4) = m^2$. Since n is square-free, we have that $n \mid (y^2 + 4)$. But $n \mid y$, so that $n \mid 4$, and hence $n = 1$ or $n = 2$. For $n = 1$, we have $y^2 + 4 = m^2$, which clearly has no positive integral solution. Letting $n = 2$, we get $8k^4 + 8 = m^2$. It follows that $8 \mid m^2$, so $m = 4t$, reducing the equation to $k^4 + 1 = 2t^2$. It is well known that all solutions of $u^4 + v^4 = 2w^2$ are of the form $u^2 = v^2 = w$. (See, for example, Dickson, *Introduction to the Theory of Numbers*, 1929, p. 43, problem 4.) Thus $k = t = 1$, so the only solution is $y = 2$, $z = 4$.

II. *Solution by E. P. Starke, Plainfield, N.J.* If $y^3 + 4y = z^2$, then $(y^3 + 4y)^2 = y^6 + 8y^4 + 16y^2 = z^4$; also $y^6 - 8y^4 + 16y^2 = (y^3 - 4y)^2$. Subtracting these two equations, we obtain $(2y)^4 = z^4 - (y^3 - 4y)^2$. But it is known that there exists no solution in nonzero integers for $u^4 - v^4 = w^2$. (See Wright, *Theory of Numbers*, 1939, p. 104, exercise 2, for example.) Hence either

$$2y = 0 \text{ and } z^4 = (y^3 - 4y)^2 \text{ or } z = 2y \text{ and } y^3 - 4y = 0.$$

Since $y > 0$, the first alternative is untenable. From the second possibility, we find the unique solution to the problem is $y = 2$, $z = 4$.

III. *Solution by E. Trost, Technikum Winterthur, Switzerland.* We consider the more general equation

$$(1) \quad ay^3 + 4a^3b^4y = z^2, \quad y, z > 0,$$

where a, b are positive integers. If (y, z) is a solution, then the quadratic polynomial $P(t) = ay^3t^2 - z^2t + 4a^3b^4y$ has the rational zero $t = 1$. Therefore $z^4 - (2aby)^4$ must be the square of an integer. Taking into account that $y > 0$ and applying a result of Fermat, we infer that $z = 2aby$. Now we see that (1) has the unique solution $(y, z) = (2ab^2, 4a^2b^3)$. (For further applications of this method, see E. Trost, *Eine Bemerkung zur Diophantischen Analysis*, Elem. der Math. 26 (1971), 60–61.)

Solutions were submitted also by forty-one others.

Return of the Rational Tangent

E 2333 [1972, 87]. *Proposed by D. E. Penney, University of Georgia*
If k, m and n are integers, then one solution of the equation

$$\frac{\pi}{4} = k \arctan \frac{m}{n}$$

is $k = m = n = 1$. Find all others.

Comment by Andrzej Makowski, Warsaw, Poland. The given equation implies that

$$\tan \frac{\pi}{4k} = \frac{m}{n}.$$

J. M. H. Olmsted (*Rational values of trigonometric functions*, this MONTHLY 52 (1945), 507–508) proved that the only rational values of $\tan \pi r$ (r a rational number) are 0 and ± 1 . By virtue of this result $m/n = 0$ or ± 1 . But $m = 0$ is impossible, so that $k = \pm 1$. Thus the only solutions are $(k, m, n) = (1, j, j)$ or $(-1, j, -j)$ where $j \neq 0$ is an arbitrary integer.

Also solved by M. T. Bird, W. J. Blundon, David Brooks, Frederick Carty, Allen Charnow & Hwa Tang, R. M. Giuli, Michael Goldberg, M. G. Greening (Australia), M. Hirschhorn (Australia), Hans Kappus (Germany), Václav Konečný, Carolyn MacDonald, L. E. Mattics, F. G. Schmitt, Jr., R. E. Shafer, G. S. Sidhu, R. Van Meter, Charles Wexler, A. Zujus, and the proposer.

Editor's comment: This old chestnut has been around for at least fifty years. The earliest reference seems to be R. S. Underwood, *Supplementary note on the irrationality of certain trigonometric functions*, this MONTHLY 29 (1922), p. 346. (This was noted by Schmitt.) Both Schmitt and Brooks note that the result can be found in Ivan Niven, *Irrational Numbers*, Carus Monograph No. 11, Corollary 3.12, p. 41. Zujus found the result in a paper by E. A. Yasinovyi in the Russian magazine, *Mathematics in School* (1958). Charnow and Tang show the analogous result for the sine (and hence the cosine): the only rational values of the sine are the “obvious” ones. Underwood seems to be the first to have noted this also — see his *On the irrationality of certain trigonometric functions*, this MONTHLY 28 (1921), 374–376.

It is interesting that even though the values of the trigonometric functions are only very rarely rational, the tabulated values are always algebraic! That is, if x is expressed as an integral number of degrees, minutes and seconds (so that it is commensurable with π) then necessarily all of the trigonometric functions of x are algebraic numbers. This was noted by Elijah Swift, *Note on trigonometric functions*, this MONTHLY 29 (1922), 404–405, and again by R. W. Hamming, *The transcendental character of $\cos x$* , this MONTHLY 52 (1945), 336–337. A more difficult question involves the degree (as an algebraic number) of the values of the trigonometric functions. See D. H. Lehmer, *A note on trigonometric algebraic numbers*, this MONTHLY 40 (1933), 165–166.

For other related problems see B. H. Arnold and H. Eves, this MONTHLY 56 (1949), 20–21, Problem 195 [1915, 27], Problem 3733 [1937, 113], and a note by Underwood in the “Question and Discussion” section [1922, 255]. Still another related problem was found by Sidhu in R. D. Carmichael, *Diophantine Analysis*, Dover, 1915. The problem is credited to Stormer (1899) and asks for all integer solutions of

$$m \arctan (1/x) + n \arctan (1/y) = k \pi/4,$$

with k, x , and y positive. It is there claimed that the only solutions are $(k, m, n, x, y) = (1, 1, 1, 2, 3)$ or $(1, 2, -1, 2, 7)$ or $(1, 2, 1, 3, 7)$ or $(1, 4, -1, 5, 239)$.

Congruence Properties of $[r^n]$

E 2334 [1972, 87]. *Proposed by Erwin Just, Bronx Community College*

Let k be an arbitrary positive integer. Prove that there exists a non-integral real number $r > 1$ with the property that k divides $[r^n]$ for every positive integer n . (The square brackets denote the greatest integer function.)

I. *Solution by J. G. Mauldon, Amherst College.* One such number is $r = k + \frac{1}{2} + \sqrt{k^2 + \frac{1}{4}}$, which is the larger root of the quadratic equation

$$x^2 - (2k + 1)x + k = 0.$$

Note that the other root s of this equation satisfies $0 < s < 1$.

Now define $u_n = r^n + s^n$ for $n = 0, 1, 2, \dots$. It can be verified that u_n satisfies the difference equation

$$u_{n+2} = (2k + 1)u_{n+1} - ku_n$$

with initial values $u_0 = 2$ and $u_1 = 2k + 1$. An obvious induction shows that u_n is an integer and $u_n \equiv 1 \pmod{k}$ for $n \geq 1$. But $r^n = u_n - s^n$, so that $u_n - 1 < r^n < u_n$ and consequently $[r^n] = u_n - 1 \equiv 0 \pmod{k}$ as required.

II. *Solution by Richard Scoville, Duke University.* We show the following more general result: Let a_1, a_2, \dots be arbitrary integers between 0 and $k - 1$ inclusive. Then there is a nonintegral $r > 1$ satisfying

$$(1) \quad [r^n] \equiv a_n \pmod{k}$$

for every $n = 1, 2, \dots$.

To show this, we choose by induction an increasing sequence $\{x_i\}$ and a decreasing sequence $\{y_i\}$ as follows: Let $x_1 = k^2 + a_1$ and $y_1 = k^2 + a_1 + 1$. Assume that x_1, \dots, x_j and y_1, \dots, y_j have been chosen so that the following are satisfied:

- (a) $x_1 < x_2 < \dots < x_j < y_j < y_{j-1} < \dots < y_1$
- (b) $y_i^i - x_i^i = 1$ for $i = 1, 2, \dots, j$
- (c) If $x_j < r < y_j$, then $[r^i] \equiv a_i \pmod{k}$ for $i = 1, 2, \dots, j$.

(Editor's comment: Note that (b) and (c) together are equivalent to the assumption that for each integer i there exists an integer m_i such that $x_i^i = km_i + a_i$ and $y_i^i = km_i + a_i + 1$. In particular, note that x_i^i and y_i^i are always integers.)

By the Mean Value Theorem and the inductive assumption (b), we have

$$y_j^{j+1} - x_j^{j+1} = (y_j^j)^{1+1/j} - (x_j^j)^{1+1/j} \geq \left(\frac{j+1}{j}\right)x_j > x_1 \geq k^2.$$

Since $y_j^{j+1} - x_j^{j+1} > k^2 > k + 1$, there are at least k integers lying strictly between x_j^{j+1} and $y_j^{j+1} - 1$; one of them is congruent to $a_{j+1} \pmod{k}$. Let it be x_{j+1}^{j+1} and define y_{j+1}^{j+1} to be $x_{j+1}^{j+1} + 1$; then

$$x_j^{j+1} < x_{j+1}^{j+1} < y_{j+1}^{j+1} < y_j^{j+1}$$

and the inductive step is completed.

The number r which can be characterized as either the common limit of the sequences $\{x_i\}$ and $\{y_i\}$ or the sole element of the set $\bigcap_{i=1}^{\infty} [x_i, y_i]$ can now be seen to have property (1).

Also solved by P. K. Garlick, G. A. Heuer, L. E. Mattics, The 3-S Group of New York, Ruby Williams, and the proposer.

Continuous Two-to-One Functions

E 2335 [1972, 88]. *Proposed by J. P. Celenza, Bayside, N.Y.*

Does there exist a continuous function from the reals to the reals which is precisely two-to-one?

Comment by M. L. Klasi and C. A. Grimm, South Dakota School of Mines and Technology. In his published solution to E 1094 [1954, 425], Azriel Rosenfeld shows that a continuous function (real-valued with domain an interval) which takes on no value more than twice must take on some value exactly once. A continuous two-to-one function therefore cannot exist.

Also solved by the proposer and 64 others.

Editor's comment: All of the solutions used in one way or another the intermediate-value property of continuous functions and the connectedness of the domain. Note however, that the transformation $f(z) = z^2$ in the complex plane maps the unit circle (which is connected) onto itself in precisely two-to-one fashion.

Reference was made to a number of related articles including O. G. Harrold, *Exactly (k , 1) transformations on connected linear graphs*, Amer. J. Math. 62 (1940), 823-834; O. G. Harrold, *The non-existence of a certain type of continuous transformation*, Duke Math. J. 5 (1939), 789-793; J. Mioduszewski, *On two-to-one continuous functions*, Rozp. Mat. 24 (1961) p. 36, and J. H. Roberts, *Two-to-one transformations*, Duke Math. J. 6 (1940), 256-262. Harrold shows that if f is a (continuous) map from an arc to an arc which is at most two-to-one and which preserves end-points, then f is necessarily a homeomorphism (and thus one-to-one). Roberts shows that there does not exist a continuous two-to-one transformation on a closed two-cell.

Generalizations are always welcomed by the editors. Several readers show that there exists a continuous n -to-one function from the reals to the reals if and only if n is odd. These readers are: David Kelly, Robert Patenaude, Howard Penn, Mary Powderly, Kenneth Rosen, E. F. Schmeichel, Gary Sherman, Robert Spira, and Clifford Wagner. Basilios Krikeles, an undergraduate, noted this theorem, but did not prove it. The non-existence of continuous even-to-one functions is proved in an analogous manner to the special case $n = 2$, whereas the existence of continuous odd-to-one functions is shown by construction. One of the simplest is due to Schmeichel: Let n be odd and define the graph of f on the interval $[k, k + 1]$ where k is any integer, to consist of the polygonal arc which joins in order the points (k, k) , $(k + 1/n, k + 1)$, $(k + 2/n, k)$, $(k + 3/n, k + 1)$, ..., $(k + 1, k + 1)$.

We note the similarity of this problem to E 1715 [1965, 784] which exhibits a continuous function on $[0, 1]$ which is precisely \aleph_0 -to-one; the result of this problem is also found in B. R. Wenner, *Continuous, exactly k -to-one functions on R* , Math. Mag., 45 (1972) 224-225.

Möbius Transformations of Finite Order

E 2336 [1972, 88]. *Proposed by William Fortney, Dumaguete City, Philippines, and Robert Breusch, Amherst College*

Consider the group of bijective rational functions over the complex numbers (with ∞) under the operation of composition. For any positive integer n , characterize the elements of order n .

Solution by J. G. Mauldon, Amherst College. Denote the given group by G ; it can be shown that $f \in G$ if and only if f is a fractional linear or Möbius transformation:

$$f(z) = \frac{az + b}{cz + d} \quad ad - bc \neq 0.$$

(A Möbius transformation is *normalized* if $ad - bc = 1$; every Möbius transformation can be normalized by dividing through all of the parameters by a square root of $ad - bc$.)

Suppose that $f \in G$ is of finite order $n > 1$. It is well known that every Möbius transformation (other than the identity) has either one or two fixed points (finite or infinite). See H. Behnke and F. Sommer, *Theorie der analytischen Funktionen einer komplexen Veränderlichen*, Springer-Verlag, Berlin, 1962, p. 327. If f has one fixed point, then f is conjugate (in the group-theoretic sense) to a transformation with a single fixed point at ∞ , i.e., to a translation. But a (nontrivial) translation cannot have finite order, and conjugate elements have the same order, so this case is impossible.

Thus f must have two distinct fixed points. If these fixed points are 0 and ∞ , then $f(z) = \alpha z$ for some fixed complex number α ; that is, f is a dilation/rotation. Since f has order n , it follows that $\alpha = \omega$, where ω is a primitive n th root of unity. In general, f is conjugate to a dilation/rotation; that is, f is conjugate to g where $g(z) = \omega z$ and ω is a primitive n th root of unity (*ibid.* pp. 327 ff.); thus f is *elliptic*. This characterizes the elements of order n : they are conjugate in G to rotations by $2\pi k/n$, where $(k, n) = 1$.

In terms of the original parameters a, b, c, d , it is known that if

$$f(z) = \frac{az + b}{cz + d}, \quad g(z) = \frac{a'z + b'}{c'z + d'},$$

where $ad - bc = a'd' - b'c' = 1$, then f and g are conjugate if and only if $a + d = \pm(a' + d')$ (*ibid.* pp. 329–330). If $g(z) = \omega z$, where $\omega = e^{2\pi i k/n}$, then we can take $a' = e^{\pi i k/n}$, $d' = e^{-\pi i k/n}$, and $b' = c' = 0$. Thus if

$$f(z) = \frac{az + b}{cz + d} \quad ad - bc = 1,$$

then f is of order $n > 1$ if and only if

$$a + d = \pm(a' + d') = \pm 2 \cos \left(\frac{\pi k}{n} \right),$$

where $(k, n) = 1$. In general (i.e., if f is not normalized), then let $\sqrt{ad - bc}$ denote either of the (complex) square roots of $ad - bc$. Then f is of order n if and only if

$$a + d = \pm 2\sqrt{ad - bc} \cos \left(\frac{\pi k}{n} \right) \quad (k, n) = 1.$$

Also solved by the proposers.

Editor's comment: Let M denote the multiplicative group of complex 2×2 matrices with determinant 1 (the *unimodular group*). Then M and G are isomorphic under

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \rightarrow f(z) = \frac{az + b}{cz + d},$$

so that the problem is essentially that of determining primitive n th roots of the identity matrix in M . This was the approach taken by the proposers in their solution. Note that $a + d$ is the trace of the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

An interesting discussion of the geometry of Möbius transformations can be found in J. Harkness and F. Morley, *Introduction to the Theory of Analytic Functions*, Macmillan, 1898, pp. 27–45 and 57–66. The authors also take up the special cases $n = 2$ and $n = 3$ of our problem.

A related problem is E 2186 [1970, 531].

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers — The State University, New Brunswick, N.J., 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before April 30, 1973. Contributors (in the United States) who desire acknowledgement of receipt of their solutions are asked to enclose self-addressed, stamped postcards.

An asterisk () means neither the proposer nor the editors supplied a solution.*

5889*. *Proposed by J. K. Doyle and F. A. Kuzam, Syracuse University*

Does there exist a topological group G with more than one point such that the fundamental group of G is isomorphic with G ?

5890*. *Proposed by Harry Ruderman, Hunter College High School*

Prove or disprove that the minimum of $|a^4 + b^4 - c^4|$ is equal to 64, where $1 \leq a < b < c$.

5891. *Proposed by V. Dlab and C. M. Ringel, Carleton University, Ottawa, Canada*

Prove or disprove that the left and right lengths of every factor ring R/I coincide, where R is a local quasi-Frobenius ring and I is an ideal of R . A local quasi-Frobenius ring is an artinian local ring with a unique minimal left and a unique minimal right ideal (which necessarily coincide).

5892*. *Proposed by John Myhill, University of Leeds, England*

Let A be a subset of the plane. For $p \in A$, $\varepsilon > 0$, let $N_\varepsilon(p) \equiv \{q \in A: d(p, q) < \varepsilon\}$. A is called *uniform* if for any $p, q \in A$ and any $\varepsilon > 0$, there is an isometry of $N_\varepsilon(p)$ onto $N_\varepsilon(q)$ taking p into q . Are there any uniform, closed connected sets other than the straight line, the circle, and the entire plane (and the empty set and a single point)?

5893. *Proposed by F. K. Dashiell, Jr., University of California at Berkeley*

A Borel subset of the unit interval $I = [0, 1]$ is called *metrically dense* if its intersection with each open interval in I has positive Lebesgue measure. Call a Borel subset D of I *metrically balanced* if both D and its complement $I - D$ are metrically dense. Prove that any Borel set S in I is the symmetric difference of two metrically balanced Borel sets D_1 and D_2 , that is, $S = (D_1 - D_2) \cup (D_2 - D_1)$.

5894. *Proposed by J. F. Kemp, Jr., Amoco Research, Tulsa, Okla.*

If $F_1(x_1), F_2(x_2), \dots, F_n(x_n)$ are n probability distribution functions, then prove that $\min(F_1(x_1), F_2(x_2), \dots, F_n(x_n))$ is an n -dimensional probability distribution function with marginals $F_1(x_1), F_2(x_2), \dots, F_n(x_n)$.

SOLUTIONS OF ADVANCED PROBLEMS

Set of Second Category

5814 [1971, 911]. *Proposed by A. C. Segal, University of Alabama, Birmingham*

It is known that a subset of the real line which is either first category or measure zero must have a void intersection with uncountably many Lebesgue cosets (i.e., cosets modulo the subgroup of rationals). Prove or disprove the converse.

Solution by the proposer. This converse is false. Let S denote the union of all rational translates of the Cantor set. S is first category, has measure zero, and completely fills the uncountably many cosets in which elements of the Cantor set appear. Therefore, T , the complement of S , has nonzero measure, is second category, and misses the uncountably many cosets which S fills.

Also solved by R. P. Boyer, and by J. C. Morgan II.

Groups with Center and Commutative Subgroup of Order p

5815 [1971, 912]. *Proposed by L. W. Shapiro, Howard University*

Show that there are no groups G of order p^{2n} with center and commutator subgroup of order p , where p is a prime.

Solution by Frank DeMeyer, Colorado State University. Let G be a p -group whose center and commutator subgroup both have order p . Let Z be the center of G .

Then G/Z is a p -group, so the center of G/Z is nontrivial. Let $x \in G - Z$ represent an element in the center of G/Z . For any $y \in G$, $[x, y] = xyx^{-1}y^{-1} \in Z$ and $x \notin Z$, so for some y , $[x, y]$ is a non identity element of Z . Since both Z and the commutator subgroup of G have order p they must coincide. Now Z is cyclic and G/Z is abelian, so Problem 5689 [1970, 1016] asserts $G/Z \cong H \times H$ for some abelian group H . Thus the order of G/Z is an even power of p , so the order of G is an odd power of p .

Also solved by D. Ž. Djoković, Vance Faber, M. G. Greening (Australia), W. H. Gustafson, W. M. Hill, A. A. Jagers (Netherlands), W. G. Leavitt, C. Y. Tang & H. T. Tang, Gomer Thomas, J. F. Watters (England), Mark Yu, and the proposer.

Combinatorics of Matrices with 0's and 1's

5816 [1971, 912]. *Proposed by Solomon Leader, Rutgers University*

Let P be a nonempty, finite set with p members, and Q be a finite set with q members. Let $N_k(p, q)$ be the number of binary relations of cardinality k with domain P and range Q . (Equivalently, $N_k(p, q)$ is the number of $p \times q$ matrices of 0's and 1's with exactly k entries equal to 1 and no row or column identically 0.) Compute $\sum_{k=1}^{pq} (-1)^{k-1} N_k(p, q)$.

Solution by Harry Lass, California Institute of Technology. Let A_i be the event that row i has all zeros, $i = 1, 2, \dots, p$, and let B_j be the event that column j has all zeros, $j = 1, 2, \dots, q$. It follows that

$$N_k(p, q) = \binom{pq}{k} - N\left(\bigcup_{i=1}^p A_i \bigcup_{j=1}^q B_j\right)$$

is the number of $p \times q$ matrices with exactly k entries equal to one, all other entries being zero, with no row or column identically zero. By the law of inclusion-exclusion we have

$$\begin{aligned} N_k(p, q) &= \binom{pq}{k} - p \binom{(p-1)q}{k} - q \binom{(q-1)p}{k} \\ &\quad + \binom{p}{2} \binom{(p-2)q}{k} + \binom{p}{1} \binom{q}{1} \binom{(p-1)(q-1)}{k} \\ &\quad + \binom{q}{2} \binom{(q-2)p}{k} - + \dots \end{aligned}$$

From $\sum_{k=1}^n (-1)^{k-1} \binom{n}{k} = 1$ for $n \geq 1$ it follows that

$$\begin{aligned} \sum_{k=1}^{pq} (-1)^{k-1} N_k(p, q) &= \sum_{r=0}^{p-1} (-1)^r \binom{p}{r} \sum_{s=0}^{q-1} (-1)^s \binom{q}{s} \\ &= (-1)^{p+1} (-1)^{q+1} = (-1)^{p+q}. \end{aligned}$$

Also solved by Robert Breusch, M. G. Greening (Australia), D. J. Kleitman, P. R. Stein, B. R. Toskey, and the proposer.

Laplace Transform of a Differentiable Function

5817 [1971, 912]. *Proposed by M. F. Neuts, Purdue University*

Let $f(t)$ be the characteristic function of a probability distribution $F(\cdot)$ whose $(n+1)$ st moment is finite. For all $\lambda > 0$ the integral $I(\lambda) = \int_0^\infty e^{-\lambda t} f(t) dt$ exists. Prove that

$$\lim_{\lambda \rightarrow +\infty} \frac{1}{I(\lambda)} \sum_{v=0}^n i^v \lambda^{-v-1} \mu'_v = 1,$$

where μ'_v is the v -th moment of $F(\cdot)$.

As a particular case, obtain the classical asymptotic expansion

$$I - \phi(\lambda) \sim -\frac{1}{\sqrt{2\pi}} e^{-\lambda^2/2} \sum_{v=1}^n (-1)^v \lambda^{-2v-1} (1 \cdot 3 \cdots (2v-1))$$

for the normal distribution function as $\lambda \rightarrow +\infty$.

Solution by A. A. Jagers, Twente University of Technology, Netherlands
From the given context it is clear that a stronger statement is possible,

$$\lim_{\lambda \rightarrow +\infty} \lambda^{n+1} (I(\lambda) - \sum_{v=0}^n i^v \lambda^{-v-1} \mu'_v) = 0.$$

Since F has a finite $(n+1)$ st moment, its characteristic function f can be expanded $n+1$ steps in a Maclaurin series:

$$f(t) = \int_{-\infty}^{\infty} e^{itx} dF(x) = \sum_{v=0}^n \frac{i^v t^v \mu'_v}{v!} + \frac{f^{(n+1)}(\theta t) t^{n+1}}{(n+1)!},$$

$0 \leq \theta \leq 1$, where $f^{(n+1)}$ is continuous and bounded by

$$\int_{-\infty}^{\infty} |x|^{n+1} dF(x) \equiv M < \infty.$$

Hence, in terms of Laplace transforms,

$$I(\lambda) = \sum_{v=0}^n i^v \lambda^{-v-1} \mu'_v + R_n(\lambda),$$

where $|R_n(\lambda)| \leq M \lambda^{-n-2}$; the statement now follows. Finally the given asymptotic expansion for the normal distribution ϕ is obtained as the particular case $F = \phi$ noting that in this case $f(t) = \exp(-t^2/2)$ and so $I(\lambda) = \sqrt{2\pi} \exp(\lambda^2/2) \cdot (1 - \phi(\lambda))$, $\mu'_v = 0$ for v odd, and $\mu'_v = 1 \cdot 3 \cdot 5 \cdots (2k-1)$ for $v = 2k$.

Also solved by S. A. Book, P. W. A. Dayananda (Singapore) A. K. Gupta & L. Jensen, J. C. Hickman, Harry Lass, O. P. Lossers (Netherlands), O. G. Ruehr, P. H. Young, and the proposer.

Simultaneous Congruences

5818 [1971, 912]. *Proposed by Erwin Just, Bronx Community College*

Let $q \geq 5$ be an integer of one of the forms $6n \pm 1$. Must there exist a prime p and an integer x for which

$$x^{q-1} - x + 1 \equiv 0 \pmod{p} \text{ and } x^q \equiv 1 \pmod{p}?$$

Solution by David Spear, student, City College, New York. We prove a stronger result: Given any positive integer q other than 1, 2, 3 or 6, there exist a prime p and an integer x such that $x^{q-1} - x + 1 \equiv 0 \pmod{p}$ and $x^q \equiv 1 \pmod{p}$.

CASE I. q is odd. Let f_n denote the n th Fibonacci number ($f_{n+1} = f_n + f_{n-1}$, $f_0 = 0$, $f_1 = 1$). It suffices to let p be any odd prime divisor of $(f_{q-1} + f_{q+1})$, if there is any, and to let $x = \frac{1}{2}(p+1)(5f_{q+1} + 1)$. For let $a = f_{q+1}$. Then

$$f_{q-1} + f_{q+1} \equiv 0 \pmod{p} \Rightarrow f_{q-1} \equiv -a \pmod{p},$$

$$f_q = f_{q+1} - f_{q-1} \Rightarrow f_q \equiv 2a \pmod{p},$$

$$f_q^2 = f_{q-1}f_{q+1} + 1 \text{ (with } q \text{ odd)} \Rightarrow 5a^2 \equiv 1 \pmod{p},$$

$$2x = (p+1)(5a+1) \Rightarrow 2x \equiv 5a+1 \pmod{p}.$$

Then $4x^2 \equiv 25a^2 + 10a + 1 \equiv 5 + 10a + 1 \equiv 2(5a+1) + 4 \equiv 4x + 4 \pmod{p}$, whence $x^2 \equiv x + 1 \pmod{p}$. A simple induction yields $x^k \equiv f_k x + f_{k-1}$, $k=1, 2, 3, \dots$. Then

$$x^q \equiv f_q x + f_{q-1} \equiv 2ax - a \equiv a(5a+1) - a \equiv 5a^2 \equiv 1 \pmod{p}.$$

Now $x(x^{q-1} - x + 1) \equiv x^q - x^2 + x \equiv 1 - x^2 + x \equiv 0 \pmod{p}$, but $x \not\equiv 0$ (because $x^q \equiv 1 \pmod{p}$) implies $x^{q-1} - x + 1 \equiv 0 \pmod{p}$.

These are the desired relations, but it remains to show that $(f_{q-1} + f_{q+1})$ must have an odd prime divisor. Suppose, instead, that $f_{q-1} + f_{q+1} = 2^s$ for some integer s . Given $q \neq 1$, $q \neq 3$. Thus $q \geq 5$, which implies $f_{q-1} + f_{q+1} \geq 8$, so that $s \geq 3$. Now $2 \mid f_n$ if and only if $3 \mid n$ and $16 \mid f_n$ if and only if $12 \mid n$. Also $f_{2q} = f_q(f_{q-1} + f_{q+1})$, so $f_{2q} = 2^s f_q$. Then

$$2^s \mid f_{2q} \Rightarrow 2 \mid f_{2q} \Rightarrow 3 \mid 2q \Rightarrow 3 \mid q \Rightarrow 2 \mid f_q \Rightarrow 2^{s+1} \mid f_{2q} \Rightarrow 16 \mid f_{2q} \Rightarrow 12 \mid 2q \Rightarrow 6 \mid q \Rightarrow 2 \mid q,$$

which is a contradiction. This completes the proof for case I.

CASE II. q is divisible by 4. Let $q = 4m$, $p = 5$, $x = 3$. Then

$$3^4 \equiv 1 \pmod{5} \Rightarrow 3^{4m} \equiv 1 \pmod{5}, \text{ i.e., } x^q \equiv 1 \pmod{p},$$

from which $3(3^{4m-1} - 3 + 1) \equiv 3^{4m} - 3^2 + 3 \equiv 1 - 9 + 3 \equiv 0 \pmod{5}$. Hence $3^{4m-1} - 3 + 1 \equiv 0 \pmod{5}$, i.e., $x^{q-1} - x + 1 \equiv 0 \pmod{p}$, fulfilling the requirements.

CASE III. q is even but not divisible by 4. Then $q = 2t$ for some odd positive integer t . Given $q \neq 2$, $q \neq 6$, we have $t \neq 1$, $t \neq 3$, therefore case I applies. That is, there exist a prime p and an integer x such that $x^{t-1} - x + 1 \equiv 0 \pmod{p}$ and $x^t \equiv 1 \pmod{p}$. Then

$$x^{q-1} - x + 1 \equiv x^{2t-1} - x + 1 \equiv x^t(x^{t-1}) - x + 1 \equiv x^{t-1} - x + 1 \equiv 0 \pmod{p}$$

and further $x^q \equiv x^{2t} \equiv (x^t)^2 \equiv 1 \pmod{p}$. This completes the proof.

Also solved by Robert Breusch, Irving Gerst, A. A. Jagers (Netherlands), L. E. Mattics, P. L. Montgomery, and the proposer.

Convergence of Function Iterates

5820 [1971, 1027]. *Proposed by Julio Cano, Findlay College, Ohio*

Let K be a compact subset of the line and f a continuous function from K to K . Suppose that $x_0 \in K$ has the property that every cluster point of the sequence $\{f^n(x_0)\}$ is a fixed point of f . Show that the sequence is convergent. Show also that this result fails in two dimensional Euclidean space.

I. *Solution by S. J. Bernau, University of Texas at Austin.* Write $x_n = f^n(x_0)$. We ignore the trivial case when some x_n is fixed by f , i.e., we assume that no x_n is a cluster point of $\{x_n\}$. Let $b = \limsup x_n$, $a = \liminf x_n$, and suppose $a < b$. If there exists k such that $x_k \in [a, b]$ then, since x_k is not a cluster point of $\{x_n\}$ there is a non-empty open interval $(c, d) \subset (a, b)$ such that $x_k \in (c, d)$ and $x_n \notin (c, d)$ if $n \neq k$. We conclude that in any case there is a non-empty interval (u, v) such that $(-\infty, u)$ and (v, ∞) both contain infinitely many x_n . Hence we choose an infinite subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $x_{n_k} < u$ and $x_{n_k+1} > v$ for all k . Clearly, no cluster point y , say, of $\{x_{n_k}\}$ can be fixed by f (continuous) since $y \leq u$ and $f(y) \geq v$. Thus $a = b$ and $\{x_n\}$ converges.

II. *Solution by R. O. Davies, The University, Leicester, England.* To show that the proposition fails in the plane, use polar coordinates and define f on the annulus $\frac{1}{2} \leq r \leq 1$ by $f((r, \theta)) = ((2-r)^{-1}, \theta + 1 - r)$. Then f is continuous, and every point of the circumference $r = 1$ is fixed. When $x_0 = (\frac{1}{2}, 1)$ we find that

$$f^n(x_0) = \left(\frac{n}{n+1}, 1 + \frac{1}{2} + \cdots + \frac{1}{n} \right);$$

hence the cluster points of $\{f^n(x_0)\}$ are the points of $r = 1$ and are fixed, but the sequence is not convergent.

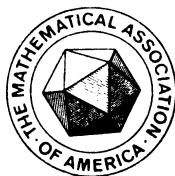
Also solved by Skagi Aggi, Max Broberg (Sweden), Bruce Ferrero, R. B. Israel, A.A. Jagers (Netherlands), J. G. Mauldon, S. S. Mitra, Nicholas Passell, and the proposer.

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF
THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 80



NUMBER 2

CODEN: AMMYAE

CONTENTS

Award for Distinguished Service to Professor Raymond L. Wilder	117
Award of the 1973 Chauvenet Prize to Professor Carl Douglas Olds	120
The Elementary Cases of Landau's Problem of Inequalities Between Derivatives	I. J. SCHOENBERG 121
Types of Fully Ordered Groups	D. P. MINASSIAN 159
The William Lowell Putnam Mathematical Competition	J. H. MCKAY 170

MATHEMATICAL NOTES

The Area of a Hypersphere in Riemannian Space	B. A. FUSARO 179
An Area Theorem for Schlicht Functions	J. L. ULLMAN 184
On Set Points of Discontinuity	O. T. ALAS 186
Generalized Fibonacci Number Triples	A. G. SHANNON AND A. F. HORADAM 187
Ambivalence in Alternating Symmetric Groups	CLAIRE PARKINSON 190

RESEARCH PROBLEMS

Can $\phi(n)$ Properly Divide $n-1$?	RONALD ALTER 192
-----------------------------------------------	------------------

CLASSROOM NOTES

A Discovery Approach to e	J. P. TULL 193
Simple Proofs of Two Estimates for e	R. B. DARST 194

MATHEMATICAL EDUCATION

The Lecture Method in Mathematics: A Student's View	M. W. HAM 195
---------------------------------------------------------------	---------------

(Continued on inside cover)

FEBRUARY

1973

ELEMENTARY PROBLEMS AND SOLUTIONS 202
ADVANCED PROBLEMS AND SOLUTIONS 208
REVIEWS 214
NEWS AND NOTICES 231
MATHEMATICAL ASSOCIATION OF AMERICA 232
Calendars of Future Meetings 232

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see Statement of Policy (vol. 76, p. 2). Manuscript preparation: Please use the Manual for Monthly Authors (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss.
Backlog: Main Articles 12 months, Math. Notes 15 months, Research Problems 7 months, Classroom Notes 11 months, Math. Education 10 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to HARLEY FLANDERS, American Mathematical Monthly, Tel Aviv University, Ramat Aviv, Israel (see Notice, vol. 77, 1970, p. 555); NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D. C. 20036.

HARLEY FLANDERS, Editor

ASSOCIATE EDITORS

- JOSHUA BARLAZ J. G. HARVEY SEYMOUR SCHUSTER
E. R. BERLEKAMP ERIC S. LANGFORD J. ARTHUR SEEBACH, Jr.
JANE W. DI PAOLA P. D. LAX E. P. STARKE
ROBERT GILMER ARTHUR MATTUCK LYNN A. STEEN
RICHARD GUY M. W. POWNALL JAMES WENDEL
RAOUL HAILPERN GIAN-CARLO ROTA

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are \$12.50. For nonmembers the subscription price is \$18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June-July, August-September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices.

Copyright © The Mathematical Association of America (Incorporated), 1973

AWARD FOR DISTINGUISHED SERVICE TO PROFESSOR RAYMOND L. WILDER

This year's Award for Distinguished Service goes to a man with wide interests. He is known for his contributions to mathematics and logic, his fondness for anthropology and the history of the development of mathematical concepts, and his service through excellence in teaching and able leadership of mathematical organizations.

Raymond L. Wilder was born in Palmer, Massachusetts in 1896. His versatility showed itself early. He was quite good at the piano, and the local movie house hired him to accompany their silent movies.

While Ray liked mathematics, logic, history, and anthropology, he did not see a ready market for a knowledge of these subjects and decided to study actuarial mathematics. Brown and Texas were leading universities in this area at that time. Wilder pursued his study of actuarial mathematics through the bachelor's and master's degree at Brown University, and then went to the University of Texas to continue these studies.

Wilder wanted a better acquaintance with pure mathematics, and asked to enroll in one of Professor R. L. Moore's classes in topology. Moore at first refused him admission, since Wilder's interest in topology was only secondary. Moore remarked that he doubted that Wilder would like the rigors of proving theorems, and perhaps would not be much good at it even if he did. However, Wilder persisted that he was really interested in pure mathematics and answered some of Moore's questions (such as, "What is an axiom?") well enough that Moore relented and let him enroll. For a while he was ignored by Moore, but as Wilder was able to prove some difficult theorems, Moore's enthusiasm began to grow. Wilder continued his actuarial studies, but when Moore learned that Wilder had solved a problem that had baffled J. R. Kline and others, he suggested that Wilder write it up for a thesis promptly. The deadline for taking language and qualifying exams had already passed, but Moore cut through red tape and arranged for Wilder to take his exams after the deadline. Wilder took his Ph. D. in topology that very same year and gave up his actuarial studies.

Wilder spent two years at Ohio State University after finishing at Texas. He then went to the University of Michigan and worked his way through the ranks to become their first Research Professor of Mathematics. Professor Wilder was very influential in helping build the University of Michigan into the giant center of excellence in mathematics that it is today. One year he was their Henry Russel Lecturer. This is the University of Michigan's highest award, presented for outstanding scholarship and to a person who has reached full maturity in his scholarly work.

Professor Wilder's students report that he is a great teacher. Not that his lectures were always polished (which they were not) but, more important, they were stimulating



RAYMOND L. WILDER

and challenging. Word got around that Wilder not only taught mathematics, but about mathematics. His course in Foundations was widely discussed, even outside mathematical circles, and the course became one of the most popular on campus. The general student learned the history of mathematical concepts and that mathematics was an important part of our culture. Each found in the course something related to his own interests. The course appealed to a diversified audience.

Professor Wilder taught all students in his classes, and not just the elite. He regarded an ordinary student reaching to the limits of his capacity as more exciting than an excellent student just coasting along. While he would not try to make a purse out of a sow's ear or turn each student into a mathematics major, he challenged each student according to that student's ability, and encouraged the student to develop his own talents. Some students who had not intended to major in mathematics changed their majors when they became impressed with the cultural importance of mathematics and were fired with the excitement of discovery. Wilder stimulated superior undergraduates to become mathematicians by admitting them to his advanced courses and letting them talk, prove theorems, and learn by doing. The article on the Distinguished Service Award to Ed Begle four years ago reported that Begle received his predoctoral training at the University of Michigan, where under the influence of Ray Wilder he became interested in topology.

Professor Wilder is superb in leading others to do research. He directed the theses of twenty-five students, including Leon Cohen, Paul Swingle, Sam Kaplan, Morton Curtis, Alice Dickinson, Joe Shoenfield, Tom Brahana, Frank Raymond, and Kyung Kwun. He influenced the work of students who wrote their theses with others, including Norman Steenrod (whose first paper came from work he did in Wilder's class) and Stephen Smale (whose first two papers were connected with Wilder's seminars and questions raised by Wilder). Wilder's seminars were lively affairs that had a stimulating impact on research. They influenced not only the work of graduate students, but also the research of Wilder himself and that of his many colleagues who attended.

One of the things that makes Professor Wilder so successful as a teacher is the high quality of his research. His earlier research dealt with properties of the plane and continuous curves, but he broadened his interests to higher dimensions and along with his students developed the important notion of generalized manifolds. He wrote over 70 research papers and three books. His colloquium volume is a profound treatment of the topology of generalized manifolds; his second book dealt with foundations; the third is concerned with the evolution of mathematical concepts. The breadth of Wilder's scholarship is evidenced by the fact that although most of his papers deal with geometric topology, he addressed the International Congress of Mathematicians on "The Cultural Basis of Mathematics," his Gibbs lecture was entitled "Trends and Social Implications of Research," and one of his latest papers appears in a medical journal.

Another quality that makes Ray Wilder a great teacher is his interest in people. The humanitarian aspect of his character is one of his finest assets. He likes people and

they like him. Students sought his council. He became to many a personal friend and one whom they never forgot. His colleagues report that students from long ago have dropped by to visit with Ray and enjoy the memories of his gentle introduction to mathematics, culture, history, and humanity that they recall with so much relish. Ray would remember them and usually prove it with an anecdote. Wilder's students have a warm place in their hearts for him and his talented and charming wife, Una.

A dramatic shift in the position of the United States in mathematical affairs occurred just prior to and during World War II. Immigration of leading foreign mathematicians played a big part in this. Professor Wilder was instrumental in influencing Michigan and the federal establishment to find places for some of these talented immigrants. His motivation was both humanitarian and scientific. Reactions such as "surely there must be an American just as well qualified" had to be overcome. Wilder's efforts in getting Sammie Eilenberg to join Michigan's staff brought great mathematical rewards. Our country today is much stronger mathematically because of the efforts of Wilder and others during these trying times.

Professor Wilder has devoted himself to working with mathematical organizations as well as to teaching and research. He has been an advisor to the Mathematics Division of the Air Force Office of Scientific Research, the NRC Fulbright Committee, the Michigan Mathematical Journal, and to many educational groups. He has been active on committees of both MAA and AMS, and has served as president of both organizations. He continues to work through these organizations for the improvement of teaching, the promotion of research, and the well-being of people.

Professor Wilder contends that mathematics is one of the most important cultural components of every modern society. He feels that a knowledge of mathematics and its methods should be a part of the intellectual and cultural background of each well-trained person; whether he be a teacher, businessman, legislator, public servant, or housewife. People well trained in mathematics are often able to analyze in a unique fashion. Wilder believes that the need for mathematically trained people is increasing — but in areas not traditionally associated with mathematics — positions in government, industry, economics, and particularly in managerial positions. New courses and new teaching techniques will be needed to give these sorts of people the special kind of mathematical training that will appeal to them and be most useful to them. We need more Wilder-type teaching.

R. H. BING

AWARD OF THE 1973 CHAUVENET PRIZE TO PROFESSOR CARL DOUGLAS OLDS

The Board of Governors of the Mathematical Association of America at its meeting on August 27, 1972, at Dartmouth College, voted to award the 1973 Chauvenet Prize to Professor Carl Douglas Olds for his paper "The Simple Continued Fraction Expansion of e ," which appeared in this MONTHLY, 77 (1970) 968-974.

A certificate and monetary award in the amount of five hundred dollars was presented to Professor Olds at the time of the annual business meeting of the Association on January 28, 1973, in Dallas.

The Chauvenet Prize is awarded for a noteworthy paper of an expository or survey nature published in English, which comes within the range of profitable reading for members of the Association. The purpose of the prize is to stimulate the writing of expository and survey articles. The 1973 Prize, awarded for a paper published in the three-year period 1969-71, is the twenty-first award of the Chauvenet Prize since its institution by the MAA in 1925. For the list of the names of previous winners, see this MONTHLY, 71 (1964) p. 589, 73(1965) pp. 2-3, 74(1967) p.3., 75(1968) pp. 3-4, 77(1970) pp. 117-118, 78(1971) pp. 112-113, and 79(1972) pp. 112-113.

Professor Olds was born on May 11, 1912, in Wanganui, New Zealand. He received all his degrees at Stanford University, the A. B. 1936, the A. M. 1937, and the Ph.D. in 1943 under Professor J. V. Uspensky. From 1935 to 1940 and in the summer of 1942, he was an acting instructor at Stanford University, and from 1940 to 1945 an assistant professor at Purdue University. Since 1945, he has been at California State University, San Jose, advancing through the ranks to full professor.

Professor Olds has served the mathematical community extensively. His service to the Mathematical Association of America included the acting chairmanship of the Northern California Section for part of 1951, Secretary-Treasurer of that Section from 1952 to 1955, and Sectional Governor for the period 1956-58. He served as first editor of the MATHEMATICAL LOG, the official publication of Mu Alpha Theta, the national high school and junior college mathematics club. He was awarded the Mu Alpha Theta service plaque in 1966.

Professor Olds' skill as a teacher was recognized by the award to him of a California State College Distinguished Teaching Award for the academic year 1965-66.

Professor Olds' substantial contributions to various parts of number theory are contained in his many publications in a great variety of periodicals. He is also the author of the book *Continued Fractions*, published as part of the New Mathematics Library of Random House in 1963.

In accepting the Award, Professor Olds indicated how very pleased and honored he felt. He added that in the past, and especially during the last few years, the Editors of the MONTHLY have done a fine job in encouraging expository writing. He thought that more mathematicians would write such articles if they realized that expository articles do not have to be long, do not have to cover an entire field of study, and do not need to include every reference that exists on a subject.

THE ELEMENTARY CASES OF LANDAU'S PROBLEM OF INEQUALITIES BETWEEN DERIVATIVES

I. J. SCHOENBERG, University of Wisconsin

INTRODUCTION

In 1913 Landau initiated in [5] a new kind of extremum problem: The sharp inequalities between the supremum-norms of derivatives. He wrote two further papers, [6] and [7], on this subject (see also [3, 139–142]). Here we are only concerned with his first paper [5]. A lively activity on this subject culminated in 1939 with Kolmogorov's remarkable paper [4], where Landau's \mathbb{R} -problem was solved for all values of n (Landau had solved it for $n = 2$ only). In 1941 Bang [2] gave a second proof of Kolmogorov's theorem using the theory of almost periodic functions. Recently, the author gave a third proof in [13]. This third proof is in essence an elaboration of Landau's original direct approach and may be regarded as an application of spline theory. The analogue of Kolmogorov's theorem for the halfline \mathbb{R}_+ has recently been established in [11].

The present paper discusses for both \mathbb{R} and \mathbb{R}_+ those cases of Landau's problem that require no knowledge beyond the elements of the Differential and Integral Calculus of functions of one variable. The novel contribution of this paper, besides the proofs, is the discussion of the extremizing functions in Theorems 4, 5, 6, and 7, for the \mathbb{R} -problem, and Theorems 9 and 11 for the \mathbb{R}_+ -problem.

The author believes that the subject can be used to supplement the contents of a calculus course, of an introductory course in numerical analysis, or for lectures in undergraduate, or beginning graduate, seminars. In doing this there is a good deal of flexibility. The main object of discussion are the Euler splines $\mathcal{E}_n(x)$, and the essential section of Part I is §2. The §§1 and 3 only furnish further background and may be omitted. If I were to make a selection, I would choose Theorems 1, 2, 4, Corollaries 1, 2, and Theorem 5. This choice was implemented on when in the

I. J. Schoenberg received his Doctor's Degree at the Univ. of Jassy, Rumania. In his thesis he initiated the theory of non-uniform asymptotic distribution of sequences, mod 1. He held positions at the Univ. of Jassy, Univ. of Chicago (Rockefeller Fellow), the Institute for Advanced Study, Swarthmore Coll., Colby Coll., and the Univ. of Pennsylvania before going to his present Professorship at the Mathematics Research Center, Univ. of Wisconsin. He has spent leaves-of-absence at the Ballistic Research Laboratories Aberdeen, Institute for Numerical Analysis U. C. L. A., Stanford Univ., I. A. S., and the Technion, Haifa.

His main research interests are Diophantine approximations, moment problems and related topics, distance geometry, total positivity, approximation theory and practice. He edited *Approximations with Special Emphasis on Spline Functions* (Academic Press 1969), and is preparing a monograph on Cardinal Spline Interpolation. *Editor.*

framework of the Visiting Lectureship Program of the MAA the author gave three one-hour lectures on this subject at Wichita State University on December 6 and 7, 1971. He wishes to thank Professor Keith Moore, Albion College, and Professor William M. Perel, Wichita State University, for arranging these lectures. This experience encouraged the author to write this paper.

I. THE EULER SPLINES

1. Cardinal spline interpolation. Let n be a natural number and let $\mathcal{S}_n = \{S(x)\}$ be the class of functions $S(x)$ having the following two properties

- (i) $S(x) \in C^{n-1}(\mathbb{R})$.
- (ii) The restriction of $S(x)$ to every interval $(v, v+1)$ between consecutive integers in a polynomial of degree $\leq n$.

Such functions $S(x)$ are called *cardinal spline function of degree n* . Evidently $\pi_n \subset \mathcal{S}_n$, where π_n denotes the class of polynomials of degree not exceeding n . We may even consider \mathcal{S}_0 , the class of step-functions with discontinuities at the integers. Indefinite integration of the elements of \mathcal{S}_0 gives the elements of \mathcal{S}_1 , also called *cardinal linear splines* (the term "spline" can be used either as an adjective or as a noun). Integrating the elements of \mathcal{S}_1 we obtain those of \mathcal{S}_2 , also called *cardinal quadratic splines a.s.f.* The term "cardinal" is to remind us that we pass from one polynomial component of $S(x)$ to the next at the integers. These transition points are called the *knots* of the spline.

It is also useful to introduce the class

$$(1.1) \quad \mathcal{S}_n^* = \{S(x); S(x + \tfrac{1}{2}) \in \mathcal{S}_n\}.$$

The elements of \mathcal{S}_n^* are again defined by the properties (i) and (ii), provided that we replace in (ii) the interval $(v, v+1)$ by $(v - \frac{1}{2}, v + \frac{1}{2})$. The knots of $S(x)$ are now half-way between the integers, and $S(x)$ may be called a *midpoint spline*.

With elements of the class \mathcal{S}_n , or perhaps \mathcal{S}_n^* , we may attempt to solve the following

CARDINAL INTERPOLATION PROBLEM. *Given the sequence of numbers*

$$(1.2) \quad (y_v) = (\dots, y_{-2}, y_{-1}, y_0, y_1, y_2, \dots)$$

we are to find $S(x)$ such that

$$(1.3) \quad S(v) = y_v \text{ for all integers } v.$$

We restrict our discussion to the case when (y_v) is a *bounded* sequence. This means that for an appropriate K

$$(1.4) \quad |y_v| < K \text{ for all } v.$$

A main result is the following

THEOREM OF CARDINAL SPLINE INTERPOLATION. *We assume that (1.4) holds.*

1. *If n is odd, then there exists a unique $S(x) \in \mathcal{S}_n$ such that $S(x)$ is bounded for all real x and satisfies the interpolation conditions (1.3).*

2. *If n is even, then there exists a unique $S(x) \in \mathcal{S}_n^*$ such that $S(x)$ is bounded and satisfies (1.3).*

The first part of this theorem was first established by Subbotin [14]. For the complete theorem under more general conditions (the condition (1.4) is replaced by the requirement that y_v should grow at most like some power of $|v|$ as $v \rightarrow +\infty$ or $v \rightarrow -\infty$) see [10].

The theorem is trivial if $n = 1$, but is no longer so if $n > 1$. Indeed, a linear spline $S_1(x)$ satisfying (1.3) is immediately obtained by successive linear interpolation between consecutive ordinates y_v and y_{v+1} . The condition (1.4) is not needed in this case and $S_1(x)$ is evidently unique for any sequence (y_v) .

Remarkable cardinal splines are obtained from the above theorem for particular simple sequences (1.2). Here are two examples.

A. *The fundamental splines.* For the special sequence

$$(1.5) \quad y_0 = 1, \quad y_v = 0 \text{ if } v \neq 0.$$

The theorem furnishes a unique bounded solution that we denote by $L_n(x)$. Thus

$$(1.6) \quad L_n(0) = 1, \quad L_n(v) = 0 \text{ if } v \neq 0.$$

Of course

$$(1.7) \quad L_n(x) \in \begin{cases} \mathcal{S}_n & \text{if } n \text{ is odd,} \\ \mathcal{S}_n^* & \text{if } n \text{ is even.} \end{cases}$$

The following is also true: The unique bounded solution $S(x)$ of the interpolation problem (1.3) may be represented by the formula

$$(1.8) \quad S(x) = \sum_{v=-\infty}^{\infty} y_v L_n(x-v),$$

where the series converges uniformly in every finite interval. This is a cardinal spline analogue of Lagrange's interpolation formula (see [10]).

B. *The Euler splines.* Very likely the most interesting examples of cardinal spline functions arise if we apply the above theorem to the sequence

$$(1.9) \quad y_v = (-1)^v \text{ for all } v.$$

For each n we denote the solution by $\mathcal{E}_n(x)$ and call it the *Euler spline of degree n* . Thus

$$(1.10) \quad \mathcal{E}_n(v) = (-1)^v \text{ for all } v, \text{ and } \mathcal{E}_n(x) \in \begin{cases} \mathcal{S}_n & \text{if } n \text{ is odd,} \\ \mathcal{S}_n^* & \text{if } n \text{ is even.} \end{cases}$$

These properties, together with the requirement that $\mathcal{E}_n(x)$ is bounded, defines this function uniquely on the basis of the cardinal interpolation theorem. We may also apply (1.8) and define $\mathcal{E}_n(x)$ by

$$\mathcal{E}_n(x) = \sum_{-\infty}^{\infty} (-1)^v L_n(x - v).$$

Our entire discussion so far was to show how the Euler splines fit into the theory of cardinal spline interpolation. However, this approach to $\mathcal{E}_n(x)$ does not help us much, because we have not established here the general interpolation theorem, nor have we learnt anything concerning $L_n(x)$ beyond its existence and uniqueness. Fortunately, there is a direct constructive approach to the Euler spline $\mathcal{E}_n(x)$ to which we now proceed.

2. A direct construction of the Euler splines. Let $f(x)$ be defined on \mathbb{R} and integrable in every finite interval.

DEFINITIONS. 1. We say that $f(x)$ is even about the point $x = a$, provided that it satisfies $f(x) = f(2a - x)$ for all x . Likewise $f(x)$ is odd about $x = a$ if $f(x) = -f(2a - x)$.

2. We say that $f(x)$ has the property P_0 , or $f(x) \in P_0$, provided that $f(x)$ is even about $x = 0$, and odd about $x = 1/2$.

3. We say that $f(x)$ has the property P_1 , or $f(x) \in P_1$, provided that $f(x)$ is odd about $x = 0$, and even about $x = 1/2$.

LEMMA 1. If $f(x) \in P_0$, or $f(x) \in P_1$, then $f(x)$ is a periodic function of period 2, hence $f(x - 2) = f(x)$.

Proof. If $f(x) \in P_0$, then

$$f(x) = -f(1 - x) = -f(x - 1) = f(2 - x) = f(x - 2).$$

If $f(x) \in P_1$, then

$$f(x) = f(1 - x) = -f(x - 1) = -f(2 - x) = f(x - 2). \quad \square$$

We may omit the proof of the easily established

LEMMA 2. If $f(x)$ is even (odd) about $x = a$ then $\int_a^x f(t)dt$ is odd (even) about $x = a$.

LEMMA 3. 1. If $f(x) \in P_0$ and $g_0(x) = \int_0^x f(t)dt$, then $g_0(x) \in P_1$.

2. If $f(x) \in P_1$ and $g_1(x) = \int_{1/2}^x f(t)dt$, then $g_1(x) \in P_0$.

Proof: 1. Let $f(x) \in P_0$. By Lemma 2 $g_0(x)$ is *odd* about $x = 0$. Let us show that it is *even* about $x = 1/2$. By Lemma 2 applied with $a = 1/2$ we have

$$\begin{aligned} g_0(x) &= \int_0^x f(t)dt = \int_0^{1/2} f(t)dt + \int_{1/2}^x f(t)dt = \int_0^{1/2} f(t)dt + \int_{1/2}^{1-x} f(t)dt \\ &= \int_0^{1-x} f(t)dt = g_0(1-x). \end{aligned}$$

2. Let $f(x) \in P_1$. By Lemma 2 $g_1(x)$ is *odd* about $x = 1/2$. Let us show that it is *even* about $x = 0$. Again, by Lemma 2

$$\begin{aligned} g_1(x) &= \int_{1/2}^x f(t)dt = \int_{1/2}^0 f(t)dt + \int_0^x f(t)dt = \int_{1/2}^0 f(t)dt + \int_0^{-x} f(t)dt \\ &= \int_{1/2}^{-x} f(t)dt = g(-x). \quad \square \end{aligned}$$

We start with the function $f_0(x)$ defined by

$$(2.1) \quad f_0(x) = (-1)^v \text{ if } v \leq x < v+1,$$

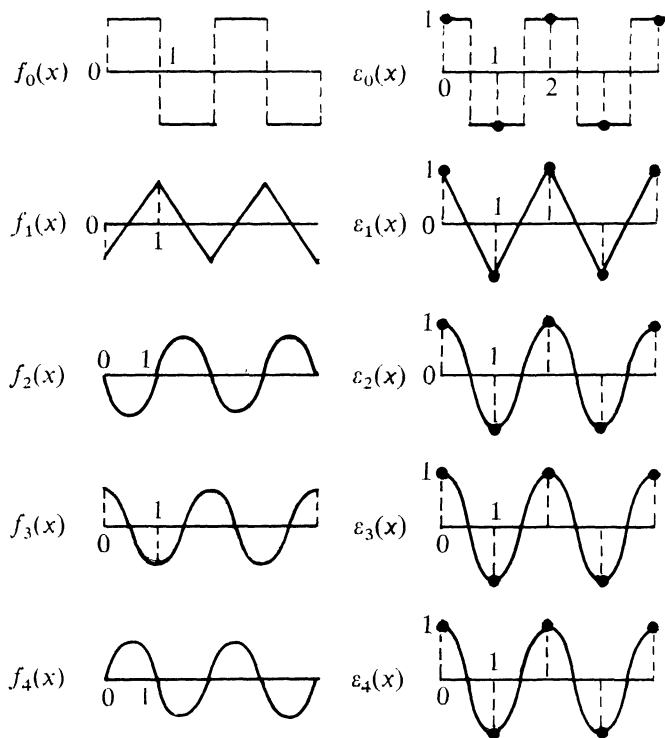


Fig. 1

whose graph is the "square-wave" of Figure 1. From it we derive the functions (see Figure 1)

$$(2.2) \quad f_1(x) = \int_{1/2}^x f_0(t)dt, f_2(x) = \int_0^x f_1(t)dt, f_3(x) = \int_{1/2}^x f_2(t)dt,$$

and generally

$$(2.3) \quad f_n(x) = \int_{\alpha_n}^x f_{n-1}(t)dt,$$

where

$$(2.4) \quad \alpha_n = \begin{cases} 0 & \text{if } n \text{ is even,} \\ 1/2 & \text{if } n \text{ is odd.} \end{cases}$$

LEMMA 4. *We have that*

$$(2.5) \quad f_n(x) \in \mathcal{S}_n, \quad (n = 0, 1, 2, \dots),$$

and

$$(2.6) \quad f_n(x) \in \begin{cases} P_0 & \text{if } n \text{ is odd,} \\ P_1 & \text{if } n \text{ is even.} \end{cases}$$

Proof: (2.5) is clear from (2.3) and an earlier remark that an integral of a spline is again a spline of a degree by one unit higher.

Also (2.6) follows from (2.3) and Lemma 3. Since $f_0(x) \in P_1$, we conclude that $f_1(x) \in P_0$ and therefore $f_2(x) \in P_1$ a.s.f. \square

LEMMA 5.1. *In $[0, 1]$ the functions $f_{2k}(x)$ are alternately strictly, convex or concave and vanish only at $x = 0$ and $x = 1$.*

2. *In $[0, 1]$ the functions $f_{2k-1}(x)$ are alternately strictly increasing or decreasing and vanish at $x = 1/2$ only.*

In particular

$$(2.7) \quad (-1)^k f_{2k-1}(0) > 0, \quad (-1)^k f_{2k}\left(\frac{1}{2}\right) > 0.$$

Proof: That $f_{2k}(x)$ vanishes at 0 follows from (2.3), (2.4), and its vanishing at 1 follows from (2.6), it being even about $x = 1/2$. (2.6) also implies that $f_{2k-1}(1/2) = 0$. The remaining statements follow from (2.3) by induction in n : $f_1(x)$ is strictly increasing, therefore $f_2(x)$ is strictly convex and therefore $f_3(x)$ is strictly decreasing. This implies the strict concavity of $f_4(x)$, a.s.f. \square

LEMMA 6. *The functions defined by*

$$(2.8) \quad \mathcal{E}_{2k-1}(x) = f_{2k-1}(x)/f_{2k-1}(0)$$

and

$$(2.9) \quad \mathcal{E}_{2k}(x) = f_{2k}(x + \tfrac{1}{2})/f_{2k}(\tfrac{1}{2})$$

are identical with the Euler splines as defined in §1B (see Figure 1).

Proof: Indeed, it should be clear that the newly defined functions enjoy the properties (1.10) and that they are bounded, since $|\mathcal{E}_n(x)| \leq 1$ for all x . The unicity of the functions having these properties establishes the identity with the old definition. In any case for us (2.8) and (2.9) is the working definition of the Euler splines. \square

If $f(x)$ is a bounded function defined on \mathbb{R} , we define its *norm* by

$$(2.10) \quad \|f\| = \sup_{x \in \mathbb{R}} |f(x)|.$$

We shall be particularly concerned with the norm of $\mathcal{E}_n(x)$ and of its derivatives and write

$$(2.11) \quad \|\mathcal{E}_n^{(v)}\| = \gamma_{n,v}, \quad (v = 0, 1, \dots, n).$$

LEMMA 7.

$$(2.12) \quad \|\mathcal{E}_n^{(v)}\| = \begin{cases} |\mathcal{E}_n^{(v)}(0)| & \text{if } v \text{ is even,} \\ |\mathcal{E}_n^{(v)}(\tfrac{1}{2})| & \text{if } v \text{ is odd.} \end{cases}$$

Proof: (2.3) implies that

$$(2.13) \quad f_n^{(v)}(x) = f_{n-v}(x) \quad (v = 0, \dots, n).$$

Moreover, we easily show that

$$(2.14) \quad \|f_n\| = \begin{cases} |f_n(\tfrac{1}{2})| & \text{if } n \text{ is even,} \\ |f_n(0)| & \text{if } n \text{ is odd.} \end{cases}$$

Let $n = 2k$, and let $c = 1/f_{2k}(\tfrac{1}{2})$. By (2.9) and (2.13) we find

$$\mathcal{E}_{2k}^{(v)}(x) = c \cdot f_{2k}^{(v)}(x + \tfrac{1}{2}) = cf_{2k-v}(x + \tfrac{1}{2}).$$

By (2.14) this is seen to reach its largest absolute value at $x = 0$ if v is even, and at $x = 1/2$ if v is odd. Similarly, using (2.8), we establish (2.12) if n is odd. \square

3. The connection with the Euler polynomials. Let us denote by $P_n(x)$ the polynomial of degree n that represents the spline function $f_n(x)$ in the interval $[0, 1]$. Thus

$$(3.1) \quad f_n(x) = P_n(x) \text{ if } 0 \leq x \leq 1, \quad P_n(x) \in \pi_n.$$

Thus, from Figure 1 we find

$$P_0(x) = 1, \quad P_1(x) = x - \tfrac{1}{2}, \quad P_2(x) = \frac{x^2}{2} - \frac{x}{2}, \quad \text{a.s.f.}$$

Clearly (2.3), (2.4) imply that

$$(3.2) \quad P_n(x) = \int_{\alpha_n}^x P_{n-1}(t)dt, \quad \alpha_n = \begin{cases} 0 & \text{if } n \text{ is even,} \\ 1/2 & \text{if } n \text{ is odd.} \end{cases}$$

and therefore

$$(3.3) \quad P'_n(x) = P_{n-1}(x).$$

A sequence of polynomials, like our $P_n(x)$, that is obtained by starting from $P_0(x) = 1$ and integrating successively, is called an *Appell sequence*. Integrating successively we obtain

$$P_1(x) = x + a_1, \quad P_2(x) = \frac{x^2}{2!} + \frac{a_1}{1!} \frac{x}{1!} + \frac{a_2}{2!}, \dots,$$

the n th polynomial being

$$(3.4) \quad P_n(x) = \frac{x^n}{n!} + \frac{a_1}{1!} \frac{x^{n-1}}{(n-1)!} + \frac{a_2}{2!} \frac{x^{n-2}}{(n-2)!} + \dots + \frac{a_{n-1}}{(n-1)!} \frac{x}{1!} + \frac{a_n}{n!}.$$

Here $a_1/1!, a_2/2!, \dots$ are the successive constants of integration.

Appell has observed that the infinite string of relations (3.4) can be described by a single relation involving series of powers of z . Indeed, multiplying the power series

$$(3.5) \quad g(z) = \sum_0^{\infty} \frac{a_n}{n!} z^n \quad \text{and} \quad e^{xz} = \sum_0^{\infty} \frac{x^n}{n!} z^n$$

and using (3.4) we find that

$$(3.6) \quad g(z)e^{xz} = \sum_0^{\infty} P_n(x)z^n.$$

The left side is called the *generating function* of the polynomials $P_n(x)$.

Let us determine $g(z)$ for the particular sequence $P_n(x)$ defined by (3.1). By (3.2) we know that

$$P_{2k}(0) = 0, \quad P_{2k-1}(\tfrac{1}{2}) = 0 \quad (k = 1, 2, 3, \dots).$$

Substituting into (3.6) the two values $x = 0$ and $x = \frac{1}{2}$, we conclude that $g(z) - 1$ is an odd function of z , and that $g(z)e^{z/2}$ is an even function of z . We therefore have the identities

$$g(z) - 1 = -g(-z) + 1 \quad \text{and} \quad g(z)e^{z/2} = g(-z)e^{-z/2}.$$

Eliminating between them $g(-z)$ we obtain that

$$(3.7) \quad g(z) = \frac{2}{e^z + 1}.$$

If we write

$$(3.8) \quad E_n(x) = n!P_n(x),$$

then (3.6) becomes

$$(3.9) \quad \frac{2e^{xz}}{e^z + 1} = \sum_0^{\infty} \frac{E_n(x)}{n!} z^n.$$

This expansion shows that the $E_n(x)$ are the classical *Euler polynomials*. (See [9], [1, Chapter 23] also for further references.) Combining (3.1) and (3.8) we obtain

$$(3.10) \quad f_n(x) = E_n(x)/n! \text{ in } 0 \leq x \leq 1,$$

and therefore, by (2.8) and (2.9), that

$$(3.11) \quad \mathcal{E}_{2k-1}(x) = E_{2k-1}(x)/E_{2k-1}(0) \text{ in } 0 \leq x \leq 1,$$

$$(3.12) \quad \mathcal{E}_{2k}(x) = E_{2k}(x + \tfrac{1}{2})/E_{2k}(\tfrac{1}{2}) \text{ in } -\tfrac{1}{2} \leq x \leq \tfrac{1}{2}.$$

The author could trace the spline function $n!f_n(x)$ to Nörlund's book [9, §16] where it is denoted by $\bar{E}_n(x)$, and where there are references to much earlier work by Hermite and Sonin (1896).

In concluding this section we mention the relations

$$(3.13) \quad \lim_{n \rightarrow \infty} L_n(x) = \frac{\sin \pi x}{\pi x}$$

and

$$(3.14) \quad \lim_{n \rightarrow \infty} \mathcal{E}_n(x) = \cos \pi x,$$

both of which hold uniformly for all real x . Concerning (3.13) see [12]. The relation (3.14) follows, via (2.8) and (2.9), from the beautiful Fourier series expansion of $f_n(x)$.

II. LANDAU'S PROBLEM FOR $\mathbb{R} = (-\infty, \infty)$. KOLMOGOROV'S THEOREM

4. Statement of Kolmogorov's theorem. Let $n \geq 2$. We consider here the class of function $f(x)$ from \mathbb{R} to \mathbb{R} that are *bounded* and have a *bounded n th derivative* $f^{(n)}(x)$. This last condition needs some further explanations as follows: In the first place we assume that

$$(4.1) \quad f(x) \in C^{n-1}(\mathbb{R})$$

and that

$$(4.2) \quad f^{(n-1)}(x) \text{ is piecewise continuously differentiable.}$$

We interpret (4.2) to mean that the graph of $f^{(n-1)}(x)$ has a continuously turning

tangent, except for corners with finite slopes for their right and left tangents, and that every finite interval contains at most a finite number of such corners. Finally, of course, $f^{(n)}(x)$ is to be bounded for all real x .

Evidently, the Euler spline $\mathcal{E}_n(x)$ satisfies all these conditions. In fact we have already considered the norms (2.11) of its derivatives and Lemma 7 shows how to identify, by (2.12), the values of

$$(4.3) \quad \gamma_{n,v} = \|\mathcal{E}_n^{(v)}\|, \quad (v = 0, 1, \dots, n), \quad \gamma_{n,0} = 1.$$

THEOREM OF KOLMOGOROV. *If $f(x)$ is such that*

$$(4.4) \quad \|f\| \leq 1, \quad \|f^{(n)}\| \leq \gamma_{n,n}$$

then

$$(4.5) \quad \|f^{(v)}\| \leq \gamma_{n,v} \text{ for } v = 1, 2, \dots, n-1.$$

The constants $\gamma_{n,v}$ in (4.5) are *best* constants because the Euler spline $\mathcal{E}_n(x)$ satisfies (4.4) and furnishes the equality sign in (4.5), simultaneously for all values of v . Complete proofs of this theorem, in the order of their appearance, are found in [4], [2], and [13]. As the title of this paper indicates we shall establish here only the cases $n = 2$, $n = 3$, and will indicate the general method of attack used in [13] by remarks concerning the problem for $n = 4$, $v = 1$, in §10.

In order to formulate the special cases that are to be established, we need the numerical values of the corresponding $\gamma_{n,v}$. From (2.8), (2.9), or by determining $f_2(x)$, $f_3(x)$, $f_4(x)$ directly by successive integrations from (2.3), we obtain

$$(4.6) \quad \mathcal{E}_2(x) = 1 - 4x^2 \text{ in } \left[-\frac{1}{2}, \frac{1}{2}\right],$$

$$(4.7) \quad \mathcal{E}_3(x) = 1 - 6x^2 + 4x^3 \text{ in } [0, 1],$$

$$(4.8) \quad \mathcal{E}_4(x) = 1 - \frac{24}{5}x^2 + \frac{16}{5}x^4 \text{ in } \left[-\frac{1}{2}, \frac{1}{2}\right].$$

Using (2.11) and (2.12), we find that

$$(4.9) \quad \gamma_{2,0} = 1, \quad \gamma_{2,1} = 4, \quad \gamma_{2,2} = 8,$$

$$(4.10) \quad \gamma_{3,0} = 1, \quad \gamma_{3,1} = 3, \quad \gamma_{3,2} = 12, \quad \gamma_{3,3} = 24,$$

$$(4.11) \quad \gamma_{4,0} = 1, \quad \gamma_{4,1} = \frac{16}{5}, \quad \gamma_{4,2} = \frac{48}{5}, \quad \gamma_{4,3} = \frac{192}{5}, \quad \gamma_{4,4} = \frac{384}{5}.$$

The first three cases of Kolmogorov's theorem may now be spelled out as follows.

THEOREM 1 (Landau). *If $f(x)$ is such that*

$$(4.12) \quad \|f\| \leq 1, \quad \|f''\| \leq 8,$$

then

$$(4.13) \quad \|f'\| \leq 4.$$

THEOREM 2. (G. E. Šilov). *If $f(x)$ is such that*

$$(4.14) \quad \|f\| \leq 1, \|f'''\| \leq 24,$$

then

$$(4.15) \quad \|f'\| \leq 3, \|f''\| \leq 12.$$

THEOREM 3. (G. E. Šilov). *If $f(x)$ is such that*

$$(4.16) \quad \|f\| \leq 1, \|f^{(4)}\| \leq \frac{384}{5},$$

then

$$(4.17) \quad \|f'\| \leq \frac{16}{5}, \|f''\| \leq \frac{48}{5}, \|f'''\| \leq \frac{192}{5}.$$

For a reference to Šilov's work see [4].

5. A kinematic interpretation: 1. It seems suggestive to think of x as time and of $f = f(x)$ as describing the motion of a point on the f -axis. The first inequality (4.12) means that the point f is forever moving on the segment $-1 \leq f \leq 1$. The second inequality (4.12) requires that the acceleration in absolute value should never exceed 8 cm/(sec)^2 . The conclusion (4.13) states that the velocity will never exceed 4 cm/sec . We know that this value is reached for the motion $f = \mathcal{E}_2(x)$ which is periodic of period 2 cm (Figure 1). Likewise (4.14) means that the rate of change of the acceleration in absolute value is not to exceed 24 cm/(sec)^3 . The conclusions concerning the velocity and acceleration are then described by the inequalities (4.15).

2. Let us consider the simple harmonic motion

$$(5.1) \quad f = \sin \omega x, \quad (\omega \text{ positive constant}).$$

By differentiation we find that

$$(5.2) \quad \|f\| = 1, \|f'\| = \omega, \|f''\| = \omega^2, \|f'''\| = \omega^3.$$

We enforce (4.12) in the most advantageous way by choosing ω such that $\omega^2 = 8$, hence $\omega = 2\sqrt{2} = 2.83$. Thus $\|f''\| = 8$, while $\|f'\| = \omega = 2.83$ falls short of the optimal value 4 given by (4.13).

Assuming (4.14) and choosing $\omega^3 = 24$, hence $\omega = 2\sqrt[3]{3} = 2.88$ we find from (5.2) that $\|f'\| = \omega = 2.88$, $\|f''\| = \omega^2 = 8.29$, which are short of the optimal values 3 and 12, respectively, as given by (4.15).

6. A general formulation of Kolmogorov's theorem. Let $F(x)$ be a bounded function

having a bounded n th derivative and let

$$(6.1) \quad \|F\| = M_0, \quad \|F^{(n)}\| = M_n.$$

What upper bound can we find for

$$(6.2) \quad \|F^{(\nu)}\| = M_\nu, \quad (0 < \nu < n)?$$

The best bound for M_ν is easily found as follows: Let a and b be positive constants and let

$$(6.3) \quad f(x) = aF(bx).$$

We shall now determine a and b such that $f(x)$ satisfies the conditions

$$(6.4) \quad \|f\| = 1, \quad \|f^{(n)}\| = \gamma_{n,n}.$$

Differentiating (6.3) and using (6.1) and (6.2), we find that

$$(6.5) \quad \|f\| = aM_0, \quad \|f^{(\nu)}\| = ab^\nu M_\nu, \quad \|f^{(n)}\| = ab^n M_n.$$

To insure (6.4) we determine a and b from the equations $aM_0 = 1$ and $ab^n M_n = \gamma_{n,n}$ and find the values

$$(6.6) \quad a = M_0^{-1}, \quad b = \gamma_{n,n}^{1/n} M_0^{1/n} M_n^{-1/n}.$$

For these values

$$(6.7) \quad \|f^{(\nu)}\| = ab^\nu M_\nu = M_0^{-1} \gamma_{n,n}^{\nu/n} M_0^{\nu/n} M_n^{-\nu/n} M_\nu.$$

The relations (6.4) show that $f(x)$ satisfies the assumptions (4.4) of Kolmogorov's theorem. We may therefore apply its conclusion to the effect that $\|f^{(\nu)}\| \leq \gamma_{n,\nu}$. Using (6.7) we find that the following statement holds.

KOLMOGOROV'S GENERAL THEOREM. *The suprema (6.1) and (6.2) satisfy the inequality*

$$(6.8) \quad M_\nu \leq C_{n,\nu} \cdot M_0^{1-(\nu/n)} M_n^{\nu/n}, \quad \text{where } C_{n,\nu} = \gamma_{n,\nu} \gamma_{n,n}^{-\nu/n} \quad (0 < \nu < n).$$

Notice that the factor $C_{n,\nu}$ is a numerical constant depending on n and ν , and that it is the best constant because we obtain equality in (6.8) for the function $F(x)$ obtained from (6.3) if we set there $f(x) = \mathcal{E}_n(x)$. This function is

$$F(x) = a^{-1} \mathcal{E}_n(b^{-1}x),$$

where a and b have the values (6.6).

Using the values (4.9) and (4.10), the inequalities (6.8) become

$$\text{for } n = 2: M_1 \leq 2^{1/2} M_0^{1/2} M_2^{1/2}$$

and

$$\text{for } n = 3: M_1 \leq (2^{-1} 3^{2/3}) M_0^{2/3} M_3^{1/3}, \quad M_2 \leq 3^{1/3} M_0^{1/3} M_3^{2/3}.$$

7. A few approximate differentiation formulae. Our immediate objective is to establish Theorems 1 and 2. For this purpose we assemble here a few simple tools.

LEMMA 8. *The following identities hold for functions $f(x)$ having appropriate derivatives which are integrable:*

$$(A) \quad f'(\tfrac{1}{2}) = f(1) - f(0) + \int_0^1 K_1(x) f''(x) dx,$$

where

$$(A') \quad K_1(x) = \begin{cases} x & \text{if } 0 \leq x \leq \tfrac{1}{2}, \\ x-1 & \text{if } \tfrac{1}{2} < x \leq 1. \end{cases}$$

$$(B) \quad f'(\tfrac{1}{2}) = f(1) - f(0) + \int_0^1 K_2(x) f'''(x) dx,$$

where

$$(B') \quad K_2(x) = \begin{cases} -\tfrac{1}{2}x^2 & \text{if } 0 \leq x \leq \tfrac{1}{2} \\ -\tfrac{1}{2}(x-1)^2 & \text{if } \tfrac{1}{2} < x \leq 1. \end{cases}$$

$$(C) \quad f''(0) = f(1) - 2f(0) + f(-1) + \int_{-1}^1 K_3(x) f'''(x) dx,$$

where

$$(C') \quad K_3(x) = \begin{cases} \tfrac{1}{2}(x+1)^2 & \text{if } -1 \leq x \leq 0 \\ -\tfrac{1}{2}(x-1)^2 & \text{if } 0 < x \leq 1. \end{cases}$$

Proof: These formulae belong to those elementary parts of numerical analysis which deal with the approximate performance of the operations of Calculus (interpolation, differentiation, integration, a.s.f.). The fundamental tool in this field is Taylor's formula with Cauchy's integral remainder

$$(7.1) \quad f(t) = f(a) + (t-a)f'(a) + \cdots + \frac{(t-a)^{n-1}}{(n-1)!} f^{(n-1)}(a) \\ + \frac{1}{(n-1)!} \int_a^t (t-x)^{n-1} f^{(n)}(x) dx.$$

It is derived by integrating by parts the remainder n times.

A. We apply (7.1) for $n = 2$, $a = 1/2$ and the two values $t = 1$ and $t = 0$, obtaining

$$f(1) = f(\tfrac{1}{2}) + \tfrac{1}{2}f'(\tfrac{1}{2}) + \int_{1/2}^1 (1-x)f''(x) dx \\ f(0) = f(\tfrac{1}{2}) - \tfrac{1}{2}f'(\tfrac{1}{2}) + \int_{1/2}^0 (-x)f''(x) dx.$$

Subtracting we get

$$f(1) - f(0) = f'(\frac{1}{2}) - \int_0^{1/2} x f''(x) dx - \int_{1/2}^1 (x-1) f''(x) dx$$

and this agrees with (A), (A').

B. Observe that $K_2(x)$ is *continuous* and that

$$K_2'(x) = -K_1(x).$$

We may therefore integrate by parts the remainder of (A) to obtain

$$\int_0^1 K_1(x) f''(x) dx = - \int_0^1 f''(x) dK_2(x) = \int_0^1 K_2(x) f'''(x) dx,$$

because $K_2(0) = K_2(1) = 0$. This establishes (B) and (B'). Alternatively, we apply (7.1) for $n = 3$, $a = 1/2$ and the two values $t = 1$ and $t = 0$, and subtract the resulting relations.

C. Apply (7.1) for $n = 3$, $a = 0$ and the two values $t = 1$ and $t = -1$ to obtain

$$f(1) = f(0) + f'(0) + \frac{1}{2} f''(0) + \frac{1}{2} \int_0^1 (1-x)^2 f'''(x) dx,$$

$$f(-1) = f(0) - f'(0) + \frac{1}{2} f''(0) + \frac{1}{2} \int_0^{-1} (-1-x)^2 f'''(x) dx.$$

Adding these we get

$$f(1) - 2f(0) + f(-1) = f''(0) - \frac{1}{2} \int_{-1}^0 (x+1)^2 f'''(x) dx + \frac{1}{2} \int_0^1 (x-1)^2 f'''(x) dx$$

which is identical with (C) and (C'). \square

8. Proofs of Theorems 1 and 2 and their extremizing functions in the strict sense.

Let us establish Theorem 1 (§4): We consider the function

$$(8.1) \quad f_0(x) = -\mathcal{E}_2(x).$$

From (4.6) and Figure 1 we see that it has the properties

$$(8.2) \quad f_0(0) = -1, f_0(1) = 1, f_0'(\frac{1}{2}) = 4 \text{ and } f_0(x) = \begin{cases} 8 & \text{in } (0, \frac{1}{2}) \\ -8 & \text{in } (\frac{1}{2}, 1). \end{cases}$$

Applying the differentiation formula (A) of §7 to $f_0(x)$ we find by (8.2) and the explicit form (A') of the kernel $K_1(x)$ that

$$(8.3) \quad 4 = f_0'(\frac{1}{2}) = 1 + 1 + 8 \int_0^1 |K_1(x)| dx.$$

Let $f(x)$ be any function satisfying (4.12) and let us evaluate $f'(\frac{1}{2})$ by the for-

mula (A). Moreover, we may assume that $f'(\frac{1}{2}) \geq 0$, for if $f'(\frac{1}{2}) < 0$ then we could replace $f(x)$ by $-f(x)$. We now obtain

$$(8.4) \quad (0 \leq) f'(\frac{1}{2}) = f(1) - f(0) + \int_0^1 K_1(x) f''(x) dx \leq 1 + 1 + 8 \int_0^1 |K_1(x)| dx$$

the last inequality being a consequence of (4.12). Moreover, the last member is equal to 4 by (8.3). Therefore

$$(8.5) \quad |f'(\frac{1}{2})| \leq 4.$$

This implies that $|f'(x_0)| \leq 4$ no matter what x_0 may be. For also $f(x + x_0 - \frac{1}{2})$ satisfies all assumptions and applying (8.5) to it, we find that $|f'(x_0)| \leq 4$. \square

Let us assume now that

$$(8.6) \quad f'(\frac{1}{2}) = 4$$

and see what the consequence are. Evidently (8.6) holds if and only if we have the equality sign in (8.4). Also, again in view of the conditions (4.12), we have equality in (8.4) if and only if $f(x)$ satisfies the conditions

$$(8.7) \quad f(0) = -1, f(1) = 1, f''(x) = \begin{cases} 8 & \text{in } (0, \frac{1}{2}) \\ -8 & \text{in } (\frac{1}{2}, 1). \end{cases}$$

Moreover, $f(0) = -1$ and $\|f\| \leq 1$ imply that $f'(0) = 0$ and $f(1) = 1$, with $\|f\| \leq 1$, imply that $f'(1) = 0$. It clearly follows from (8.6) that

$$f(x) = -\mathcal{E}_2(x) \text{ in } (0, 1).$$

We state this result as

THEOREM 4. *If*

$$(8.8) \quad \|f\| \leq 1, \|f''\| \leq 8$$

and

$$(8.9) \quad f'(\frac{1}{2}) = 4,$$

then

$$(8.10) \quad f(x) = -\mathcal{E}_2(x) \text{ in the interval } [0, 1].$$

Outside the interval $[0, 1]$ there is little that we can say about the function $f(x)$ satisfying (8.8) and (8.9). Indeed, notice that there are many ways in which the function (8.10) can be extended to all reals and still satisfying (8.8) (of course with the *equality* sign in both inequalities). For beside the obvious extension

$$(8.11) \quad f(x) = -\mathcal{E}_2(x) \text{ for all real } x,$$

we can also write

$$(8.12) \quad f(x) = \begin{cases} 1 & \text{if } x > 1 \\ -\mathcal{E}_2(x) & \text{if } 0 \leq x \leq 1, \\ -1 & \text{if } x < 0, \end{cases}$$

and many similar modifications of the function (8.11).

A comment on the function $f(x)$ satisfying (8.8) and (8.9) is in order. We can call $f(x)$ an *extremizing function* in Theorem 1 because $f(x)$ satisfies (4.13) with the equality sign, hence

$$(8.13) \quad \|f'\| = 4.$$

Moreover, we wish to call this $f(x)$ an extremizing function *in the strong sense* because the supremum of $|f'(x)|$ ($= 4$) is *actually assumed for a real x , viz. $x = \frac{1}{2}$* . We shall see in §9 that there are numerous extremizing function $f(x)$ in Theorem 1, hence satisfying (8.13), such that

$$(8.14) \quad |f'(x)| < 4 \text{ for all real } x.$$

Such functions may be called *extremizing functions in the weak sense*.

Let us establish Theorem 2 (§4): Let $f(x)$ satisfy (4.14) $\|f\| \leq 1$, $\|f'''\| \leq 24$, and let us show that (4.15) $\|f'\| \leq 3$, $\|f''\| \leq 12$.

We reproduce here the second differentiation formula

$$(B) \quad f'(\tfrac{1}{2}) = f(1) - f(0) + \int_0^1 K_2(x)f'''(x)dx$$

of Lemma 8 and apply it to the function

$$(8.15) \quad f_0(x) = -\mathcal{E}_3(x) = -1 + 6x^2 - 4x^3 \text{ in } [0, 1].$$

This function has in $[0, 1]$ the properties

$$(8.16) \quad f_0(0) = -1, f_0(1) = 1, f_0'''(x) = -24.$$

In view of (B'), of Lemma 8, we know that $K_2(x) < 0$ in $(0, 1)$, and from (B) we derive

$$(8.17) \quad 3 = f_0'(\tfrac{1}{2}) = 1 + 1 + 24 \int_0^1 |K_2(x)| dx.$$

If $f(x)$ is any function satisfying (4.14), let us evaluate $f'(\frac{1}{2})$ by (B), assuming that $f'(\frac{1}{2}) \geq 0$ (otherwise we take $-f(x)$). We obtain

$$(8.18) \quad (0 \leq) f'(\tfrac{1}{2}) = f(1) - f(0) + \int_0^1 K_2(x)f'''(x)dx \leq 1 + 1 + 24 \int_0^1 |K_2(x)| dx = 3,$$

by (8.17) and the first inequality (4.15) is thereby established.

At this point we interrupt our proof of Theorem 2 in order to see what we can say about $f(x)$ if

$$(8.19) \quad f'(\tfrac{1}{2}) = 3,$$

i.e., if equality holds in (8.18). From (4.14) we see that we have equality in (8.18) if and only if $f(x)$ has the properties

$$(8.20) \quad f(0) = -1, f(1) = 1, \text{ and } f'''(x) = -24 \text{ in } (0, 1).$$

However, as before, we also have

$$(8.21) \quad f'(0) = f'(1) = 0$$

and, of course, (8.19). The conditions (8.19), (8.20) and (8.21) are more than sufficient to imply

$$(8.22) \quad f(x) = -\mathcal{E}_3(x) \text{ in } [0, 1].$$

Let us record here this result as

COROLLARY 1. *If $f(x)$ satisfies (4.14) and (8.19), then (8.22) also holds.*

We now wish to establish the second inequality (4.15): For this purpose we need the third formula

$$(C) \quad f''(0) = f(1) - 2f(0) + f(-1) + \int_{-1}^1 K_3(x)f'''(x)dx$$

of Lemma 8. We recall that by the formula (C') of that lemma the kernel has the properties

$$(8.23) \quad K_3(x) > 0 \text{ in } (-1, 0), K_3(x) < 0 \text{ in } (0, 1).$$

We now apply (C) to the function

$$(8.24) \quad f_0(x) = -\mathcal{E}_3(x) = \begin{cases} -1 + 6x^2 + 4x^3 & \text{in } [-1, 0] \\ -1 + 6x^2 - 4x^3 & \text{in } (0, 1]. \end{cases}$$

This function has the properties

$$(8.25) \quad f_0(-1) = 1, f_0(0) = -1, f_0(1) = 1, f_0'''(x) = \begin{cases} 24 & \text{in } (-1, 0), \\ -24 & \text{in } (0, 1), \end{cases}$$

and (C), (8.23), and (8.25) show that

$$(8.26) \quad 12 = f_0''(0) = 1 + 2 + 1 + 24 \int_{-1}^1 |K_3(x)| dx.$$

If $f(x)$ is any function satisfying (4.14), and assuming that $f''(0) \geq 0$, an application

of (C) shows that

$$\begin{aligned}
 (0 \leq) f''(0) &= f(1) - 2f(0) + f(-1) + \int_{-1}^1 K_3(x) f'''(x) dx \\
 (8.27) \qquad &\leq 1 + 2 + 1 + 24 \int_{-1}^1 |K_3(x)| dx = 12
 \end{aligned}$$

by (8.26). Applying this result to $f(x + x_0)$ we obtain that $|f''(x_0)| \leq 12$, and Theorem 2 is established. \square

Let us assume that $f(x)$, satisfying (4.14), is such that

$$(8.28) \qquad f''(0) = 12,$$

and let us examine the consequences of this assumption. Clearly (8.28) if and only if we have the equality sign in (8.27) and this turn holds if and only if

$$f(-1) = 1, f(0) = -1, f(1) = 1, \text{ and } f'''(x) = \begin{cases} 24 & \text{in } (-1, 0), \\ -24 & \text{in } (0, 1). \end{cases}$$

From this we conclude that

$$(8.29) \qquad f(x) = -\mathcal{E}_3(x) \text{ in } -1 \leq x \leq 1.$$

We have therefore established the

COROLLARY 2. *If $f(x)$ satisfies (4.14) and (8.28) holds, then also (8.29) holds.*

The following generalization follows by a change of origin:

COROLLARY 2'. *If $f(x)$ satisfies (4.14) and is such that*

$$(8.30) \qquad f''(a) = \pm 12$$

then

$$(8.31) \qquad f(x) = \mp \mathcal{E}_3(x - a) \text{ if } a - 1 \leq x \leq a + 1.$$

We may now state our

THEOREM 5. *If (4.14) $\|f\| \leq 1$, $\|f'''\| \leq 24$ and if in one of the inequalities (4.15) $\|f'\| \leq 3$, $\|f''\| \leq 12$, we have the equality sign, the corresponding supremum being actually attained, then*

$$(8.32) \qquad f(x) = \mathcal{E}_3(x - c) \text{ for all real } x,$$

for an appropriate constant c .

Proof: 1. Let us assume that $\|f'\| = 3$. This supremum being assumed, we loose no generality by assuming that

$$(8.33) \quad f'(\tfrac{1}{2}) = 3.$$

Now Corollary 1 implies that

$$(8.34) \quad f(x) = -\mathcal{E}_3(x) \text{ in } [0, 1].$$

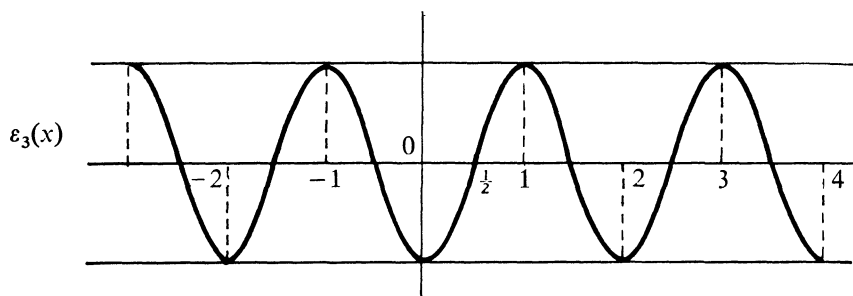


FIG. 2.

This in turn shows that $f''(1) = -12$ (see Figure 2) and now Corollary 2' shows $f(x) = -\mathcal{E}_3(x)$ in $[0, 2]$. But then surely $f''(2) = 12$ and Corollary 2' implies that $f(x) = -\mathcal{E}_3(x)$ in $[1, 3]$. We can continue in this way indefinitely and conclude that $f(x) = -\mathcal{E}_3(x)$ for $x \geq 0$. However, the same reasoning works also to the left: From (8.34) we conclude that $f''(0) = +12$ and therefore (8.34) holds also in $[-1, 1]$, hence $f''(-1) = -12$ and (8.34) holds in $[-2, 0]$ a.s.f. Therefore $f(x) = -\mathcal{E}_3(x) = \mathcal{E}_3(x-1)$ holds for all real x .

2. If we have equality in the second inequality (4.15), we get the same conclusion by applying only Corollary 2'. \square

9. The extremizing functions in the weak sense. In the present section we discuss only the cubic case of $n = 3$. Our last Theorem 5 has answered the question as to when we have the equality sign in one of the inequalities (4.15) for the case when the respective supremum is actually attained.

DEFINITION 4. We say that $f(x)$ is an extremizing function in the weak sense for $n = 3$, provided that $f(x)$ satisfies the inequalities

$$(9.1) \quad \|f\| \leq 1, \quad \|f'''\| \leq 24,$$

and therefore also

$$(9.2) \quad \|f'\| \leq 3, \quad \|f''\| \leq 12,$$

with the equality sign in one of the inequalities (9.2), the corresponding supremum not being attained.

This definition raises the following questions:

QUESTION 1. Do extremizing functions in the weak sense exist?

QUESTION 2. Let us suppose that they do and let $f(x)$ be one such. Does then the equality sign hold in all four inequalities (9.1), (9.2)?

We shall see that the answers to both questions are *affirmative*.

The affirmative answer to Question 1 is contained in

THEOREM 6. *There exist functions $f(x)$ such that*

$$(9.3) \quad \|f\| = 1, \|f'\| = 3, \|f''\| = 12, \|f'''\| = 24,$$

while

$$(9.4) \quad |f(x)| < 1, |f'(x)| < 3, |f''(x)| < 12, |f'''(x)| < 24 \text{ for all real } x.$$

Proof: We know that $f(x) = \mathcal{E}_3(x)$ satisfies (9.3), but not (9.4). To enforce both (9.3) and (9.4) we let the function “sag between $-\infty$ and $+\infty$ ” by passing to the new function

$$(9.5) \quad f(x) = \mathcal{E}_3(x)\phi(x)$$

with an appropriate positive function $\phi(x)$ to be constructed.

Using the known values (4.10) of $\gamma_{3,v} = \|\mathcal{E}^{(v)}\|$ we derive from (9.5) the inequalities

$$(9.6) \quad \begin{aligned} |f(x)| &= |\mathcal{E}\phi| \leq \phi(x), \\ |f'(x)| &= |\mathcal{E}'\phi + \mathcal{E}\phi'| \leq 3\phi(x) + |\phi'(x)|, \\ |f''(x)| &= |\mathcal{E}''\phi + 2\mathcal{E}'\phi' + \mathcal{E}\phi''| \leq 12\phi(x) + 6|\phi'(x)| + |\phi''(x)|, \\ |f'''(x)| &= |\mathcal{E}'''\phi + 3\mathcal{E}''\phi' + 3\mathcal{E}'\phi'' + \mathcal{E}\phi'''| \leq 24\phi(x) + 36|\phi'(x)| \\ &\quad + 9|\phi''(x)| + |\phi'''(x)|. \end{aligned}$$

We shall therefore satisfy (9.4) if $\phi(x)$ is *positive* and such that

$$\begin{aligned} \phi &< 1, \\ 3\phi + |\phi'| &< 3, \\ 12\phi + 6|\phi'| + |\phi''| &< 12, \\ 24\phi + 36|\phi'| + 9|\phi''| + |\phi'''| &< 24, \text{ for all real } x. \end{aligned}$$

These amount to

$$\begin{aligned} 1 - \phi &> 0, \\ 1 - \phi &> \frac{1}{3}|\phi'|, \\ 1 - \phi &> \frac{1}{2}|\phi'| + \frac{1}{12}|\phi''|, \\ 1 - \phi &> \frac{3}{2}|\phi'| + \frac{3}{8}|\phi''| + \frac{1}{24}|\phi'''|. \end{aligned}$$

Observe that the last inequality implies the previous ones. It suffices therefore to require that $\phi(x)$ be positive and to satisfy the differential inequality

$$(9.7) \quad 1 - \phi(x) > \frac{3}{2} |\phi'(x)| + \frac{3}{8} |\phi''(x)| + \frac{1}{24} |\phi'''(x)| \text{ for all real } x.$$

In order to insure also the equations (9.3), it is clear that $\phi(x)$ should also satisfy the boundary conditions

$$(9.8) \quad \phi(x) \rightarrow 1, \phi'(x) \rightarrow 0, \phi''(x) \rightarrow 0, \phi'''(x) \rightarrow 0 \text{ as } x \rightarrow \pm \infty.$$

Indeed, Leibniz's formulae (see (9.6)!) and the periodicity of $\mathcal{E}(x)$ will then show that

$$\|f^{(v)}\| \geq \overline{\lim}_{x \rightarrow \pm \infty} |(\mathcal{E}\phi)^{(v)}| = \|\mathcal{E}^{(v)}\|, \quad (v = 0, 1, 2, 3).$$

Let

$$(9.9) \quad \psi(x) = 1 - e^{-\gamma x}, \quad (\gamma \text{ positive constant}).$$

A simple calculation shows that $\psi(x)$ will surely satisfy (9.7), provided that

$$(9.10) \quad 1 > \frac{3}{2}\gamma + \frac{3}{8}\gamma^2 + \frac{1}{24}\gamma^3$$

for which $0 < \gamma \leq 1/2$ will certainly do. We now define

$$(9.11) \quad \phi(x) = \begin{cases} 1 - e^{-\gamma x} & \text{if } x \geq 1, \\ \phi(-x) & \text{if } x \leq -1. \end{cases}$$

Assuming (9.10), this function satisfies (9.7) outside the interval $(-1, 1)$. Moreover, $\phi(x)$ is positive and also satisfies the boundary conditions (9.8).

There remains to bridge the gap between -1 and 1 and this we do by interpolation as follows. Let

$$(9.12) \quad P(x) = A + Bx^2 + Cx^4$$

and

$$(9.13) \quad \phi(x) = P(x) \text{ in } -1 \leq x \leq 1.$$

We also require $P(x)$ to satisfy the interpolatory conditions

$$(9.14) \quad P(1) = \psi(1), P'(1) = \psi'(1), P''(1) = \psi''(1).$$

The functions $P(x)$ and $\phi(x)$ being both even, it is clear that the requirements (9.14) will insure that $\phi(x) \in C''(\mathbb{R})$.

We are yet to insure that $P(x)$ is positive and satisfies (9.7) in $[0, 1]$, and therefore also in $[-1, 1]$. From (9.14) we easily get for the coefficients of $P(x)$ the values

$$(9.15) \quad A = 1 - (1 + \frac{5}{8}\gamma + \frac{1}{8}\gamma^2)e^{-\gamma}, \quad B = \frac{1}{4}\gamma(3 + \gamma)e^{-\gamma}, \quad C = -\frac{1}{8}\gamma(1 + \gamma)e^{-\gamma}.$$

1. *The positivity of $P(x)$ in $[0, 1]$:* Dropping the positive term Bx^2

$$\begin{aligned} P(x) &= A + Bx^2 + Cx^4 > A + C \\ &= 1 - (1 + \frac{5}{8}\gamma + \frac{1}{8}\gamma^2)e^{-\gamma} - \frac{1}{8}\gamma(1 + \gamma)e^{-\gamma} > 0 \end{aligned}$$

because the last inequality is equivalent to $e^\gamma > 1 + \frac{6}{8}\gamma + \frac{2}{8}\gamma^2$, which evidently holds.

2. *$P(x)$ satisfies (9.7) in $[0, 1]$:* We are to find γ such that

$$(9.16) \quad 1 - A - Bx^2 - Cx^4 > \frac{3}{2}|2Bx + 4Cx^3| + \frac{3}{8}|2B + 12Cx^2| + \frac{1}{24}|24Cx|$$

holds in $0 \leq x \leq 1$. Dropping on the left the positive term $-Cx^4$, cancelling the common factor $e^{-\gamma}$, and taking on the left side all terms with their negative values for $x = 1$ and on the right with their positive values for $x = 1$, we easily find, after rearrangements that the inequality (9.16) is surely satisfied if the inequality $1 > \frac{1}{8}(35\gamma + 25\gamma^2)$ holds. This is the case if

$$(9.17) \quad 0 < \gamma < \frac{1}{5}.$$

To summarize: Let γ satisfy (9.17) and $P(x)$ be defined by (9.12) and (9.15). Finally let

$$\phi(x) = \begin{cases} 1 - e^{-\gamma|x|} & \text{if } |x| \geq 1 \\ P(x) & \text{if } -1 < x < 1. \end{cases}$$

Then $f(x)$, defined by (9.5), satisfies the conditions (9.3) and (9.4) of Theorem 6. \square

The second question is answered affirmatively by

THEOREM 7. *Let $f(x)$ be such that*

$$(9.18) \quad \|f\| \leq 1, \quad \|f'''\| \leq 24$$

and therefore

$$(9.19) \quad \|f'\| \leq 3, \quad \|f''\| \leq 12.$$

If the equality sign holds in one of the inequalities (9.19), then the equality sign holds in all four inequalities.

Proof: 1. Let us assume that

$$(9.20) \quad \|f'\| = 3.$$

If the supremum $\|f'\|$ is assumed, then we know by Theorem 5 that (8.32) holds

and we are through. We may therefore assume that

$$(9.21) \quad |f'(x)| < 3 \text{ for all real } x.$$

Let (x_v) , $(v = 1, 2, \dots)$, be a sequence of points such that

$$(9.22) \quad \lim_{v \rightarrow \infty} f'(x_v) = 3,$$

the reasoning to be applied being similar if this limit should be -3 . It should be clear that the sequence (x_v) can not have a finite limit point ξ , for we could then conclude from the continuity of $f'(x)$ that $f'(\xi) = 3$, in contradiction to (9.21). We may therefore assume that $x_v \rightarrow +\infty$, or perhaps $-\infty$. Let us assume that

$$(9.23) \quad \lim_{v \rightarrow \infty} x_v = +\infty.$$

By the formula (B) of Lemma 8 we may write

$$(9.24) \quad f'(x_v) = f(x_v + \tfrac{1}{2}) - f(x_v - \tfrac{1}{2}) + \int_0^1 K_2(x) f'''(x + x_v - \tfrac{1}{2}) dx,$$

while (8.17) shows that

$$(9.25) \quad 3 = 1 + 1 + \int_0^1 |K_2(x)| \cdot 24 dx.$$

From (9.22) we conclude that

$$(9.26) \quad \begin{aligned} & f(x_v + \tfrac{1}{2}) - f(x_v - \tfrac{1}{2}) + \int_0^1 |K_2(x)| \{ -f'''(x + x_v - \tfrac{1}{2}) \} dx \\ & \rightarrow 1 + 1 + \int_0^1 |K_2(x)| \cdot 24 dx \quad \text{as } v \rightarrow \infty. \end{aligned}$$

From this relation we shall derive all that we need.

We observe first that

$$f(x_v + \tfrac{1}{2}) - f(x_v - \tfrac{1}{2}) + \int_0^1 |K_2| \{ -f''' \} dx > 1 + 1 + \int_0^1 |K_2| \cdot 24 dx - \varepsilon$$

if $v > N(\varepsilon)$, while

$$\int_0^1 |K_2| 24 dx \geq \int_0^1 |K_2| \{ -f''' \} dx$$

and $1 \geq -f(x_v - \tfrac{1}{2})$ hold anyway. Adding these three inequalities we find that $f(x_v + \tfrac{1}{2}) > 1 - \varepsilon$ if $v > N(\varepsilon)$ and therefore

$$(9.27) \quad \lim_{v \rightarrow \infty} f(x_v + \tfrac{1}{2}) = 1.$$

Similarly we find that

$$(9.28) \quad \lim_{v \rightarrow \infty} f(x_v - \tfrac{1}{2}) = -1,$$

and finally, from (9.26), that

$$\lim_{v \rightarrow \infty} \int_0^1 |K_2(x)| \{-f'''(x + x_v - \tfrac{1}{2})\} dx = \int_0^1 |K_2(x)| \cdot 24 dx.$$

If we write

$$(9.29) \quad \phi_v(x) = 24 + f'''(x + x_v - \tfrac{1}{2}), \quad (0 \leq x \leq 1),$$

we know that this sequence of piece-wise continuous functions has the properties

$$(9.30) \quad 0 \leq \phi_v(x) \leq 48 \text{ in } [0, 1],$$

and

$$(9.31) \quad \lim_{v \rightarrow \infty} \int_0^1 |K_2(x)| \phi_v(x) dx = 0.$$

From (B') of §7 we know that $|K_2(x)|$ vanishes at 0 and 1, that it increases in $[0, \frac{1}{2}]$ and decreases in $[\frac{1}{2}, 1]$. Also that $K_2(x) = K_2(1 - x)$. We choose α such that $0 < \alpha < \frac{1}{2}$ and may write

$$(9.32) \quad \int_0^1 |K_2(x)| \phi_v(x) dx \geq |K_2(\alpha)| \int_\alpha^{1-\alpha} \phi_v(x) dx \geq |K_2(\alpha)| \cdot \inf_{[\alpha, 1-\alpha]} \phi_v(x).$$

Now (9.31) implies that

$$(9.33) \quad \inf_{[\alpha, 1-\alpha]} \phi_v(x) \rightarrow 0 \text{ as } v \rightarrow \infty.$$

Selecting ξ_v in $[\alpha, 1 - \alpha]$ such that $\phi_v(\xi_v) < \inf \phi_v(x) + 2^{-v}$, we conclude from (9.33) that $\phi_v(\xi_v) \rightarrow 0$. Finally, returning to f''' by (9.29) we have shown that

$$(9.34) \quad \lim_{v \rightarrow \infty} f'''(\xi_v + x_v - \tfrac{1}{2}) = -24.$$

Evidently (9.27), or (9.28), and (9.34) show that

$$(9.35) \quad \|f\| = 1, \quad \|f'''\| = 24.$$

There remains to show that

$$(9.36) \quad \|f''\| = 12.$$

From (9.31) and (9.32) we conclude that

$$\lim_{v \rightarrow \infty} \int_\alpha^{1-\alpha} \phi_v(x) dx = 0.$$

However, this integral can be evaluated by (9.29) and we obtain

$$24(1 - 2\alpha) + f''(\eta_v) - f''(\xi_v) \rightarrow 0,$$

where $\eta_v = 1 - \alpha + x_v - \frac{1}{2}$, $\xi_v = \alpha + x_v - \frac{1}{2}$. Therefore $f''(\xi_v) - f''(\eta_v) \rightarrow 24(1 - 2\alpha)$ as $v \rightarrow \infty$, hence $f''(\xi_v) - f''(\eta_v) > 24 - 48\alpha - \varepsilon$ if $v > N(\varepsilon)$. Adding to this the inequality $f''(\eta_v) \geq -12$ we obtain that $f''(\xi_v) > 12 - 48\alpha - \varepsilon$ if $v > N(\varepsilon)$. Since $\xi_v \rightarrow +\infty$ and α is arbitrarily small, we conclude that

$$\lim_{x \rightarrow +\infty} \overline{f''(x)} \geq 12.$$

This, together with $\|f''\| \leq 12$, shows that (9.36) holds.

2. A similar method, this time using the approximate differentiation formula (C) of Lemma 8, allows to show that (9.36) implies the equality sign in all other inequalities (9.18) and (9.19). However, we omit further details. \square

3. There are theorems analogous to Theorems 6 and 7 for the case when $n = 2$ and they are easier to derive. Also for $n = 2$ there are extremizing functions in the weak sense, i.e., satisfying (8.8), (8.13), and (8.14). The details may be left to the reader.

10. How is Theorem 3 established? As its title indicates, this paper is devoted to the elementary cases of Landau's problem. However, Theorem 3 is no longer an elementary case. The ideas underlying its proof are just as simple as before, but the necessary tools, i.e., the required approximate differentiation formulae, are more complicated.

Let us sketch, with a minimum of detail, a proof of the first inequality

$$(10.1) \quad \|f'\| \leq 16/5$$

of Theorem 3, assuming that

$$(10.2) \quad \|f\| \leq 1, \|f^{(4)}\| \leq 384/5.$$

The approximate differentiation formula that we need is

$$(10.3) \quad \begin{aligned} f'(\tfrac{1}{2}) = & \mu f(1) + \mu\lambda f(2) + \mu\lambda^2 f(3) + \cdots \\ & - \mu f(0) - \mu\lambda f(-1) - \mu\lambda^2 f(-2) \cdots + \int_{-\infty}^{\infty} K(x) f^{(4)}(x) dx, \end{aligned}$$

where

$$(10.4) \quad \mu = \frac{-60 + 12\sqrt{30}}{5} = 1.14534, \lambda = -11 + 2\sqrt{30} = -.045548.$$

The kernel $K(x)$ is a cardinal cubic spline, i.e., having its knots at the integers, except that at $x = \frac{1}{2}$ it has a discontinuity in its second derivative. It satisfies $K(x) = -K(1-x)$ and is therefore odd about the point $x = 1/2$. $K(x)$ decays

exponentially as $x \rightarrow \pm \infty$ so that $K(x)$ is absolutely integrable on the real axis. Moreover

$$(10.5) \quad K(v + \tfrac{1}{2}) = 0 \quad \text{for all integer } v,$$

and $K(x)$ vanishes nowhere else. Finally

$$(10.6) \quad K(x) < 0 \quad \text{if } -\tfrac{1}{2} < x < \tfrac{1}{2}$$

and it changes sign at each $v + \frac{1}{2}$ (see Figure 3).

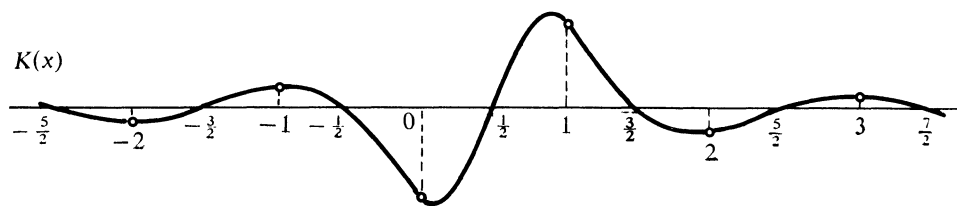


FIG. 3.

If we substitute into (10.3) the function

$$f_0(x) = -\mathcal{E}_4(x),$$

we find that $K(x)f_0^{(4)}(x)$ is positive for all x , except that it vanishes if $x = v + \frac{1}{2}$ by (10.5). Since $f_0^{(4)}(x) = \pm 384/5$ we obtain the result

$$(10.7) \quad \frac{16}{5} = f_0' \left(\frac{1}{2} \right) = 2\mu \sum_0^\infty |\lambda|^v + \frac{384}{5} \int_{-\infty}^\infty |K(x)| dx.$$

If $f(x)$ is any function satisfying (10.2), and assuming that $f'(\frac{1}{2}) \geq 0$, we obtain from (10.3) and (10.2) the estimate

$$0 \leq f' \left(\frac{1}{2} \right) \leq 2\mu \sum_0^\infty |\lambda|^v + \frac{384}{5} \int_{-\infty}^\infty |K(x)| dx = \frac{16}{5},$$

by (10.7). By reasonings used before, the equality sign is seen to hold only if $f(x) = -\mathcal{E}_4(x)$. This establishes (10.1), except that we have not proved the identity (10.3), nor do we propose to do so. However, let me say the following: The formula (10.3) is exact, i.e., its remainder vanishes, whenever $f(x)$ is a cubic polynomial. This clearly does not characterize the formula. However, (10.3) can be shown to be exact if $f(x)$ is a cardinal cubic spline with knots at $v + \frac{1}{2}$ that grows at most like a power of $|x|$ as $x \rightarrow \pm \infty$, and this condition characterizes the formula (10.3) and allows to derive it.

A last remark: The question arises whether (10.3) could be replaced in the above application by some appropriate *finite* formula that involves only finitely many of the ordinates $f(v)$. The answer is no: It can be shown that no finite differentiation formula exists that will serve the same purpose. For further details we refer to [13].

III. LANDAU'S PROBLEM FOR $\mathbb{R}_+ = [0, \infty)$.

11. The case $n = 2$. Landau's problem for the halfline \mathbb{R}_+ is similar to the problem solved by Kolmogorov's theorem, the difference being that now the competition is open only for functions from \mathbb{R}_+ to \mathbb{R} . Accordingly, the role of the previous norm $\|f\|$ is now taken over by the *halfline norm*

$$(11.1) \quad \|f\|_+ = \sup |f(x)| \text{ for } x \geq 0.$$

To facilitate the comparison with the results of Part II, we choose the same normalization as in Kolmogorov's theorem, namely

$$(11.2) \quad \|f\|_+ \leq 1, \quad \|f^{(n)}\|_+ \leq \gamma_{n,n},$$

the objective being to find within this class those functions that maximize the norms

$$(11.3) \quad \|f^{(v)}\|_+, \text{ for } v = 1, 2, \dots, n-1.$$

The transition to other normalizations, such as the one used in [11], can be achieved by means of the trivial transformation (6.3) used in §6. In §13 the \mathbb{R}_+ -analogue of Kolmogorov's theorem will be mentioned. In the meantime we turn to the first of the two elementary cases of the problem.

We assume that

$$(11.4) \quad \|f\|_+ \leq 1, \quad \|f''\|_+ \leq 8,$$

and seek a function $f_0(x)$ such that $\|f_0'\|_+ \geq \|f'\|_+$ for all functions $f(x)$ satisfying (11.4).

The function $\mathcal{E}_2(x)$ satisfies (11.4) and we also know that $\|\mathcal{E}_2'\|_+ = 4$ (this is where the constant 4 of Theorem 1 came from). Now we can do better! Indeed, let us consider $\mathcal{E}_2(x)$ for $x \geq -\frac{1}{2}$, and let us remove the knot at $x = -\frac{1}{2}$ and continue the quadratic $y = 1 - 4x^2$ (see (4.6)) also for values of $x < -\frac{1}{2}$, until we reach the point where the parabolic graph of $y = 1 - 4x^2$ intersects the horizontal line $y = -1$. We find that this happens for $x = -1/\sqrt{2}$. We consider the function

$$g(x) = \begin{cases} 1 - 4x^2 & \text{if } -1/\sqrt{2} \leq x \leq 0, \\ \mathcal{E}_2(x) & \text{if } 0 \leq x < \infty, \end{cases}$$

and shift the origin to $-1/\sqrt{2}$ to define

$$f_0(x) = g(x - 1/\sqrt{2}) \text{ for } x \geq 0, \text{ (see Figure 4).}$$

Clearly

$$(11.5) \quad \|f_0\|_+ = 1, \quad \|f_0\|_+ = 8.$$

However, it should also be clear from Figure 4 that $\|f_0'\|_+$ is reached by $|f_0'(x)|$ for $x = 0$ so that

$$\|f_0'\|_+ = f_0'(0) = g'(-1/\sqrt{2}) = -8x|_{x=-1/\sqrt{2}} = 4\sqrt{2} = 5.65684.$$

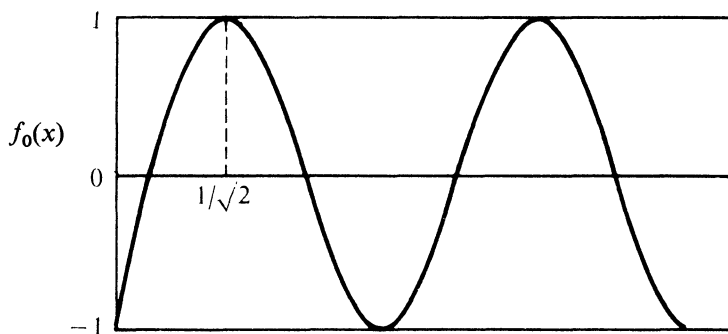


FIG. 4.

Therefore

$$(11.6) \quad \|f'_0\|_+ = f'_0(0) = 4\sqrt{2}.$$

We see that the conditions (11.4) no longer imply that $\|f'\|_+ \leq 4$, as in Theorem 1, but allow considerably larger values such as $\|f'\|_+ = 4\sqrt{2}$. This, however, is the largest value, a fact which we state as

THEOREM 8. (Landau's theorem). *If*

$$(11.7) \quad \|f\|_+ \leq 1, \quad \|f''\|_+ \leq 8,$$

then

$$(11.8) \quad \|f'\|_+ \leq 4\sqrt{2}.$$

Here $4\sqrt{2}$ is the best constant because it is reached for the above function $f_0(x)$.

Proof: As in the case of Theorem 1, we need a differentiation formula. By Taylor's formula (7.1), for $n = 2$, $a = 0$ and $t = 1/\sqrt{2}$, we have

$$f(1/\sqrt{2}) = f(0) + (1/\sqrt{2})f'(0) + \int_0^{1/\sqrt{2}} (1/\sqrt{2} - x)f''(x)dx.$$

Solving for $f'(0)$ we obtain

$$(11.9) \quad f'(0) = \sqrt{2}f(1/\sqrt{2}) - \sqrt{2}f(0) - \int_0^{1/\sqrt{2}} (1 - x\sqrt{2})f''(x)dx.$$

Applying this to $f_0(x)$ we find that

$$(11.10) \quad 4\sqrt{2} = f'_0(0) = \sqrt{2} + \sqrt{2} + 8 \int_0^{1/\sqrt{2}} (1 - x\sqrt{2})dx.$$

If $f(x)$ is any function satisfying (11.7), and assuming $f'(0) \geq 0$ (else we work with $-f$), and estimating $f'(0)$ from (11.9) and (11.7) we obtain

$$0 \leq f'(0) \leq \sqrt{2} + \sqrt{2} + 8 \int_0^{1/\sqrt{2}} (1 - x\sqrt{2})dx = 4\sqrt{2}$$

by (11.10). Therefore

$$(11.11) \quad |f'(0)| \leq 4\sqrt{2}.$$

However, if $f(x)$ satisfies (11.7), also $f(x + x_0)$, with $x_0 \geq 0$, will satisfy (11.7). We may therefore apply to $f(x + x_0)$ our previous conclusion (11.11) to infer that $|f'(x_0)| \leq 4\sqrt{2}$. Therefore $\|f'\|_+ \leq 4\sqrt{2}$. \square

We turn now to the *extremizing functions*. Unlike the situation discussed in §9 (for $n = 3$) there are no extremizing functions in the weak sense in the present case of \mathbb{R}_+ . In fact we have the following very precise theorem.

THEOREM 9. *If $f(x)$ satisfies (11.7) and if*

$$(11.12) \quad \|f'\|_+ = 4\sqrt{2}$$

then

$$(11.13) \quad f(x) = \pm f_0(x) \text{ in the interval } 0 \leq x \leq 1/\sqrt{2}.$$

REMARK. Beyond (11.13) there is little that can be said about the extremizing function $f(x)$. Indeed, the function (11.13) can be continued from $1/\sqrt{2}$ to $+\infty$ in various ways, such as $f(x) = \pm 1$ if $x > 1/\sqrt{2}$, or else by $f(x) = \pm \mathcal{E}_2(x - (1/\sqrt{2}))$, without violating the basic condition (11.7).

Proof: We distinguish two cases depending on whether the supremum $\|f'\|_+$ is attained or not.

1. Let us assume that it is attained and that

$$(11.14) \quad \hat{f}'(\xi) = 4\sqrt{2},$$

for if this value were $-4\sqrt{2}$ we could work with $-f(x)$. Let us write

$$(11.15) \quad K(x) = 1 - x\sqrt{2}, \quad (0 \leq x \leq 1/\sqrt{2})$$

for the kernel in the formula (11.9). By (11.9) and (11.10) we conclude that the equation (11.14) is equivalent to

$$(11.16) \quad \begin{aligned} \sqrt{2}f\left(\xi + \frac{1}{\sqrt{2}}\right) - \sqrt{2}f(\xi) - \int_0^{1/\sqrt{2}} K(x)f''(x + \xi)dx \\ = \sqrt{2} + \sqrt{2} + 8 \int_0^{1/\sqrt{2}} K(x)dx. \end{aligned}$$

Because $K(x)$ is positive in $[0, 1/\sqrt{2})$, (11.16) and (11.7) imply that

$$(11.17) \quad f(\xi) = -1, f\left(\xi + \frac{1}{\sqrt{2}}\right) = 1, \text{ and } f''(x + \xi) = -8 \text{ in } \left(0, \frac{1}{\sqrt{2}}\right).$$

Clearly $\xi = 0$, for if ξ were positive, then $f(\xi) = -1$ and $\|f\|_+ \leq 1$, would imply that $f'(\xi) = 0$, in contradiction to the assumption (11.14). Now (11.17) reduce to

$$f(0) = -1, f(1/\sqrt{2}) = 1, f''(x) = -8 \text{ in } (0, 1/\sqrt{2}),$$

and this already implies that $f(x) = -1 + 4\sqrt{2}x - 4x^2 = f_0(x)$ in $[0, 1/\sqrt{2}]$. Therefore (11.13) is established for this case.

2. Let us assume that

$$(11.18) \quad |f'(x)| < 4\sqrt{2} \text{ for } x \geq 0,$$

and let us show that this can not possibly happen.

Indeed, the assumption (11.12) implies the existence of an infinite sequence (x_v) of points of \mathbb{R}_+ , such that

$$(11.19) \quad \lim_{v \rightarrow \infty} f'(x_v) = 4\sqrt{2},$$

where on the right we have chosen the positive sign without loss of generality. On the other hand we have the following: If

$$(11.20) \quad x \geq \frac{1}{2},$$

then the formula (A) of Lemma 8, the relation (8.3), and the assumptions (11.7) show that

$$\begin{aligned} |f'(x)| &= \left| f(x + \tfrac{1}{2}) - f(x - \tfrac{1}{2}) + \int_0^1 K_1(t) f''(t + x - \tfrac{1}{2}) dt \right| \\ &\leq 1 + 1 + 8 \int_0^1 |K_1(t)| dt = 4. \end{aligned}$$

Thus (11.20) implies that

$$(11.21) \quad |f'(x)| \leq 4.$$

From (11.19) we now conclude (observe that $4 < 4\sqrt{2}$!) that

$$(11.22) \quad 0 \leq x_v \leq \tfrac{1}{2} \text{ for } v \text{ sufficiently large, } v > N \text{ say.}$$

From the Bolzano-Weierstrass theorem we infer that the sequence (x_v) has a limit point ξ in $[0, \frac{1}{2}]$ and therefore

$$(11.23) \quad \lim_{v' \rightarrow \infty} x_{v'} = \xi,$$

where v' is an appropriate increasing sequence of integers. The continuity of $f'(x)$ now implies that

$$4\sqrt{2} = \lim f'(x_{v'}) = f'(\lim x_{v'}) = f'(\xi).$$

Therefore $f'(\xi) = 4\sqrt{2}$, in contradiction to (11.18). The second possibility therefore never arises and Theorem 9 is established. \square

12. The case $n = 3$. As in the previous case we retain the conditions (4.14) of Theorem 2 but this time for \mathbb{R}_+ , hence

$$(12.1) \quad \|f\|_+ \leq 1, \quad \|f''\|_+ \leq 24,$$

and wish to find $f_0(x)$ satisfying (12.1) and having the largest possible value for the norm $\|f'_0\|_+$ of its first derivative. We also seek (perhaps another) $f_0(x)$ satisfying (12.1) and maximizing $\|f''_0\|_+$. We shall see that one and the same function $f_0(x)$ will do both. From (4.3) and (4.10) we know that

$$(12.2) \quad \|\mathcal{E}_3\|_+ = 1, \quad \|\mathcal{E}'_3\|_+ = 3, \quad \|\mathcal{E}''_3\|_+ = 12, \quad \|\mathcal{E}'''_3\|_+ = 24,$$

so that $\mathcal{E}_3(x)$ satisfies (12.1). However, by an appropriate modification of $\mathcal{E}_3(x)$ we can increase considerably the norms of f' and f'' .

To obtain the modified function $f_0(x)$ we remove the knot $x = 0$ of $\mathcal{E}_3(x)$, and continue its cubic polynomial branch (4.7), hence $1 - 6x^2 + 4x^3$, for negative values of x until it intersects the line $y = -1$. This happens for $x = -\frac{1}{2}$ and we define the function

$$(12.3) \quad g(x) = \begin{cases} 1 - 6x^2 + 4x^3 & \text{if } -\frac{1}{2} \leq x \leq 0, \\ \mathcal{E}_3(x) & \text{if } x \geq 0. \end{cases}$$

For technical reasons we shift the origin to the point $x = -\frac{1}{2}$ and define

$$(12.4) \quad f_0(x) = g(x - \tfrac{1}{2}) \quad (\text{see Figure 5}).$$

Notice that $f_0(x)$ is a cubic spline in \mathbb{R}_+ having no longer a knot at $x = \frac{1}{2}$, in fact

$$(12.5) \quad f_0(x) = -1 + 9x - 12x^2 + 4x^3 \quad \text{if } 0 \leq x \leq 3/2.$$

Clearly

$$(12.6) \quad \|f_0\|_+ = 1, \quad \|f'_0\|_+ = 24.$$

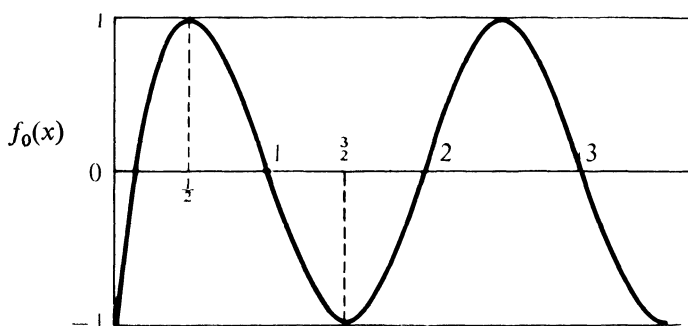


FIG. 5.

Moreover, we verify easily from (12.5) that $|f'_0(x)|$ reaches its largest value for $x = 0$, hence

$$(12.7) \quad \|f'_0\|_+ = f'_0(0) = 9.$$

Similarly we find that also $|f''_0(x)|$ reaches its largest value for $x = 0$. From (12.5) we read off this value to be

$$(12.8) \quad \|f''_0\| = -f''_0(0) = 24.$$

Comparing (12.7) and (12.8) with (4.15), we see that $f_0(x)$ surpasses by far the corresponding bounds of Theorem 2. These, however, are the largest possible values, as stated by

THEOREM 10. (A. P. Matorin). *If*

$$(12.9) \quad \|f\|_+ \leq 1, \quad \|f'''\|_+ \leq 24,$$

then

$$(12.10) \quad \|f'\|_+ \leq 9, \quad \|f''\|_+ \leq 24.$$

In (12.10) the constants 9 and 24 are the best constants because they are reached by the above function $f_0(x)$, (see [8]).

Proof: We need two differentiation formulae that we get from (7.1). Applying (7.1) for $n = 3$, $a = 0$, and the two values $t = \frac{1}{2}$ and $t = \frac{3}{2}$, we obtain

$$f\left(\frac{1}{2}\right) = f(0) + \frac{1}{2}f'(0) + \frac{1}{8}f''(0) + \frac{1}{2} \int_0^{1/2} \left(\frac{1}{2} - x\right)^2 f'''(x) dx,$$

$$f\left(\frac{3}{2}\right) = f(0) + \frac{3}{2}f'(0) + \frac{9}{8}f''(0) + \frac{1}{2} \int_0^{3/2} \left(\frac{3}{2} - x\right)^2 f'''(x) dx.$$

Solving these equations for $f'(0)$ and $f''(0)$ we obtain the two formulae

$$(12.11) \quad f'(0) = -\frac{8}{3}f(0) + 3f\left(\frac{1}{2}\right) - \frac{1}{3}f\left(\frac{3}{2}\right) + \int_0^{3/2} K_4(x)f'''(x)dx,$$

where

$$(12.12) \quad K_4(x) = \begin{cases} x\left(1 - \frac{4}{3}x\right) & \text{if } 0 \leq x \leq \frac{1}{2}, \\ \frac{1}{6}\left(\frac{3}{2} - x\right)^2 & \text{if } \frac{1}{2} < x \leq \frac{3}{2}, \end{cases}$$

and

$$(12.13) \quad f''(0) = \frac{8}{3}f(0) - 4f\left(\frac{1}{2}\right) + \frac{4}{3}f\left(\frac{3}{2}\right) + \int_0^{3/2} K_5(x)f'''(x)dx,$$

where

$$(12.14) \quad K_5(x) = \begin{cases} \frac{4}{3}\left(x^2 - \frac{3}{4}\right) & \text{if } 0 \leq x \leq \frac{1}{2}, \\ -\frac{2}{3}\left(x - \frac{3}{2}\right)^2 & \text{if } \frac{1}{2} < x \leq \frac{3}{2}. \end{cases}$$

Notice that in each of these formulae the coefficients of $f(0)$, $f\left(\frac{1}{2}\right)$, and $f\left(\frac{3}{2}\right)$ alternate in sign and that

$$(12.15) \quad K_4(x) > 0 \text{ and } K_5(x) < 0 \text{ in } 0 < x < \frac{3}{2}.$$

We now return to the function $f_0(x)$ defined by (12.4) and graphed in Figure 5. From Figure 5 and (12.5) we gather the following properties:

$$(12.16) \quad f_0(0) = -1, f_0\left(\frac{1}{2}\right) = 1, f_0\left(\frac{3}{2}\right) = -1,$$

$$(12.17) \quad f_0'(0) = 9, f_0''(0) = -24,$$

$$(12.18) \quad f_0'''(x) = 24 \text{ in } \left[0, \frac{3}{2}\right).$$

Applying the identities (12.11) and (12.13) to $f_0(x)$, we obtain by (12.15) the relations

$$(12.19) \quad f_0'(0) = 9 = \frac{8}{3} + 3 + \frac{1}{3} + 24 \int_0^{3/2} |K_4(x)| dx,$$

$$(12.20) \quad -f_0''(0) = 24 = \frac{8}{3} + 4 + \frac{4}{3} + 24 \int_0^{3/2} |K_5(x)| dx.$$

If $f(x)$ is a function satisfying the conditions (12.9), we can estimate its derivatives at the origin by (12.11) and (12.13), and obtain

$$|f'(0)| \leq \frac{8}{3} + 3 + \frac{1}{3} + 24 \int_0^{3/2} |K_4(x)| dx,$$

$$|f''(0)| \leq \frac{8}{3} + 4 + \frac{4}{3} + 24 \int_0^{3/2} |K_5(x)| dx.$$

The right hand sides being equal to 9 and 24, respectively, in view of (12.19) and (12.20), we conclude that

$$|f'(0)| \leq 9, |f''(0)| \leq 24.$$

Applying this result to $f(x + x_0)$, where $x_0 > 0$, we obtain (12.10). \square

We shall now investigate the extremizing functions in Matorin's Theorem 10

and shall see that extremizing functions in the weak sense do not exist. We begin with

LEMMA 9. 1. If $f(x)$ satisfies (12.9) and

$$(12.21) \quad |f'(\xi)| = 9 \text{ for some } \xi \geq 0,$$

then necessarily $\xi = 0$ and

$$(12.22) \quad f(x) = \pm f_0(x), \quad (x \geq 0),$$

where $f_0(x)$ is the function defined by (12.4) (Figure 5).

2. The same conclusions ($\xi = 0$ and (12.22)) hold if

$$(12.23) \quad |f''(\xi)| = 24 \text{ for some } \xi \geq 0.$$

Proof: 1. Let us first assume that $\xi = 0$ hence

$$(12.24) \quad f'(0) = 9.$$

By an oft repeated argument we conclude from (12.11) and (12.19), that (12.24) is equivalent to the relation

$$(12.25) \quad -\frac{8}{3}f(0) + 3f\left(\frac{1}{2}\right) - \frac{1}{3}f\left(\frac{3}{2}\right) + \int_0^{3/2} K_4(x)f'''(x)dx = \frac{8}{3} + 3 + \frac{1}{3} + 24 \int_0^{3/2} |K_4(x)|dx,$$

and that this implies that

$$(12.26) \quad f(0) = -1, f\left(\frac{1}{2}\right) = 1, f\left(\frac{3}{2}\right) = 1, f'''(x) = 24 \text{ in } \left(0, \frac{3}{2}\right).$$

This information already suffices to conclude that

$$(12.27) \quad f(x) = f_0(x) \text{ in } [0, 3/2].$$

But then $f''(3/2) = 12$ (see Figure 5). By Corollary 2' we now conclude that the identity (12.27) can be extended to $[1/2, 5/2]$. Continuing in this manner we see that (12.27) holds for $x \geq 0$.

Let us now show that ξ must vanish. Indeed, if

$$(12.28) \quad \xi > 0 \text{ and } f'(\xi) = 9 \text{ (say),}$$

then as above we conclude, as in (12.26), that $f(\xi) = -1$, a.s.f. But then we must have $f'(\xi) = 0$ (or else $\|f\|_+ \leq 1$ would be violated!), which contradicts the assumption $f'(\xi) = 9$.

2. If (12.23) holds, we apply similar reasonings using formulae (12.13) and (12.20). \square

THEOREM 11. *Let*

$$(12.29) \quad \|f\|_+ \leq 1, \quad \|f'''\|_+ \leq 24,$$

and therefore

$$(12.30) \quad \|f'\|_+ \leq 9, \quad \|f''\|_+ \leq 24.$$

If the equality sign holds in one of the inequalities (12.30), then

$$(12.31) \quad f(x) = \pm f_0(x) \text{ for } x \geq 0,$$

where $f_0(x)$ is the function defined by (12.4) (Figure 5).

Proof: 1. Let us suppose that

$$(12.32) \quad \|f'\|_+ = 9.$$

If this supremum is assumed, hence (12.21) holds, then the conclusion (12.31) is already assured by Lemma 9. We may therefore assume that

$$(12.33) \quad |f'(x)| < 9 \text{ for } x \geq 0,$$

and let us show that *this can not happen* by reaching a contradiction.

By (12.32) and (12.33), there exists an infinite sequence (x_v) of points of \mathbb{R}_+ such that

$$(12.34) \quad \lim_{v \rightarrow \infty} f'(x_v) = 9.$$

(If this limit were -9 we could work with $-f(x)$). In the interval $x \geq 1/2$ we can apply the differentiation formula (B) of Lemma 8, in the form

$$f'(x) = f(x + \tfrac{1}{2}) - f(x - \tfrac{1}{2}) + \int_0^1 K_2(t) f''(t + x - \tfrac{1}{2}) dt$$

to conclude from (8.17) that $|f'(x)| \leq 1 + 1 + 24 \int_0^1 |K_2(t)| dt = 3$. Thus

$$(12.35) \quad |f'(x)| \leq 3 \text{ if } x \geq \tfrac{1}{2}.$$

Confronting (12.34) with (12.35) we conclude that $0 \leq x_v \leq \frac{1}{2}$ if $v > N$. The Bolzano-Weierstrass theorem insures the existence of an appropriate infinite sequence of increasing integers (v') such that

$$(12.36) \quad \lim_{v' \rightarrow \infty} x_{v'} = \xi, \text{ for some } \xi \text{ within } [0, \tfrac{1}{2}].$$

Using the continuity of $f'(x)$, we conclude from (12.36) and (12.34) that $f'(\xi) = 9$, which contradicts our assumption (12.33).

2. If $\|f''\|_+ = 24$, we may use entirely similar arguments. If the supremum is assumed, we use Lemma 9. That the supremum is always assumed is shown by

contradiction as above: Formula (C) of Lemma 8 shows that $|f''(x)| \leq 12$ if $x \geq 1$ (here we use (8.26)!), and the continuity of $f''(x)$ takes care of the rest. \square

13. The case $n = 4$ is not elementary. Our success in attacking the Landau problem for \mathbb{R}_+ for $n = 2$ and $n = 3$ with the modified Euler splines $f_0(x)$ seems surprising, to say the least. However, for $n = 4$ this approach does not work anymore. To make it clear why, let us try to do it. Our problem is to study functions satisfying

$$(13.1) \quad \|f\|_+ \leq 1, \quad \|f^{(4)}\|_+ \leq 384/5 = 76.8$$

and to determine within this class the best, or least, constants $\gamma_{4,v}^+$ such that

$$(13.2) \quad \|f^{(v)}\|_+ \leq \gamma_{4,v}^+ \quad (v = 1, 2, 3).$$

Stated equivalently: *Within the class of functions satisfying (13.1) we wish to maximize each of the three norms on the left side of (13.2).*

We start from $\mathcal{E}_4(x)$. From (4.8) we know that

$$P(x) = 1 - \frac{24}{5}x^2 + \frac{16}{5}x^4 = \mathcal{E}_4(x) \text{ if } -\frac{1}{2} \leq x \leq \frac{1}{2}.$$

We consider $\mathcal{E}_4(x)$ for $x \geq -\frac{1}{2}$ only, and remove its knot at $x = -\frac{1}{2}$ to continue the graph of the quartic $P(x)$ for $x \leq -\frac{1}{2}$. We find that it has a minimum value at $x = -\sqrt{3}/2 = -.866$, where it assumes the value $-4/5$, and thereafter increases to $+\infty$ as $x \rightarrow -\infty$. The new function so obtained satisfies the second condition (13.1). However, to satisfy also the first condition (13.1), we must cut it off at the point where it intersects the line $y = 1$. This is found to take place at $x = -\sqrt{6}/2 = -1.225$. Accordingly we define

$$g(x) = \begin{cases} 1 - \frac{24}{5}x^2 + \frac{16}{5}x^4 & \text{in } [-\sqrt{6}/2, 0], \\ \mathcal{E}_4(x) & \text{in } [0, \infty). \end{cases}$$

As before, we shift the origin to $-\sqrt{6}/2$ and define the function

$$(13.3) \quad f_0(x) = g\left(x - \frac{\sqrt{6}}{2}\right) \text{ for } x \geq 0 \text{ (see Figure 6).}$$

We find that

$$(13.4) \quad \begin{aligned} \|f_0'\|_+ &= -f_0'(0) = 48\sqrt{6}/10 = 11.7576 \\ \|f_0''\|_+ &= f_0''(0) = 48 \\ \|f_0'''\|_+ &= -f_0'''(0) = 384\sqrt{6}/10 = 94.0604. \end{aligned}$$

These values are surely *lower bounds* for the best constants $\gamma_{4,v}^+$ of (13.2). However, our $f_0(x)$ is *certainly not an extremizing function*. This can be seen from Figure 6

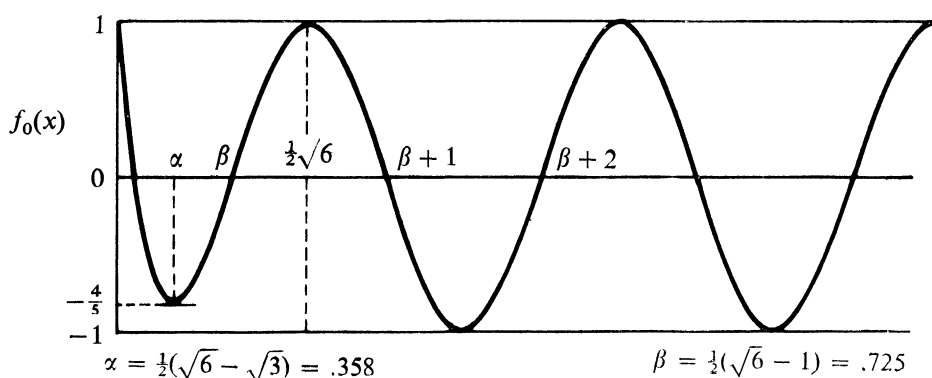


FIG. 6.

because the first minimum value of $f_0(x)$ is $= -4/5$ and thereby fails to reach down to the line $y = -1$. However, I do not know any explicitly defined function $f(x)$, satisfying (13.1), whose norms are superior to the norms (13.4) of $f_0(x)$.

At this point we state (see [11])

THE \mathbb{R}_+ -ANALOGUE OF KOLMOGOROV'S THEOREM. Let $n \geq 2$. There is a spline function $e_n(x)$ of degree n , satisfying $\|e_n\|_+ = 1$, $\|e_n^{(n)}\|_+ = \gamma_{n,n}$, with the following property: If

$$(13.5) \quad \|f\|_+ \leq 1, \quad \|f^{(n)}\|_+ \leq \gamma_{n,n},$$

then

$$(13.6) \quad \|f^{(v)}\|_+ \leq \|e_n^{(v)}\|_+ = |e_n^{(v)}(0)|, \quad (v = 1, 2, \dots, n-1).$$

These are the best constants because we have equalities if $f(x) = e_n(x)$. If $n \geq 3$, then $\pm e_n(x)$ are the only functions with these properties.

We call $e_n(x)$ the one-sided Euler spline of degree n . Just like $\mathcal{E}_n(x)$, also $e_n(x)$ has the property that $e_n^{(n)}(x)$ is a step-function assuming the values $\pm \gamma_{n,n}$ only. Figures 4 and 5 show the graphs of $e_2(x)$ and $e_3(x)$, respectively. The knots of $f_0(x)$ (Figure 6) are at its zeros $\beta + 1, \beta + 2, \dots$. The graph of $e_4(x)$ looks much like the graph of $f_0(x)$ (Figure 6), except that also its first minimum is $= -1$. However, the knots of $e_4(x)$ do not agree with its zeros, but approach them in the limit as we approach $+\infty$.

No explicit expressions are known for $e_n(x)$ ($n \geq 4$). Rather $e_n(x)$ is defined in [11] as the limit of a sequence of spline functions of degree n , that are themselves defined by minimum properties. In deriving the numerical results of [11] good approximations of $e_n(x)$, for $n = 4, 5, 6$, are used. These approximations furnish for $n = 4$ the values of the best constants in (13.2):

$$\gamma_{4,1}^+ = -e_4'(0) = 12.695$$

$$\gamma_{4,2}^+ = e_4''(0) = 50.393$$

$$\gamma_{4,3}^+ = -e_4'''(0) = 96.197.$$

In conclusion let me say the following. The Landau problems are *extremum problems*. Faced with an extremum problem we are often well on the way to its solution, provided that we are lucky enough to guess what the extremizing function is. The extremizing functions $\mathcal{E}_n(x)$ of the \mathbb{R} -problem are beautiful, simple, and easily computable functions. This is decidedly not the case of $e_n(x)$, if $n \geq 4$, and this is the reason why the \mathbb{R}_+ -problem was more difficult to solve.

This work was sponsored by the Mathematics Research Center, Madison, Wisconsin, under Contract No. DA-31-124-ARO-D-462.

References

1. M. Abramovitz and I. A. Stegun (Editors), Handbook of mathematical functions with formulas, graphs and mathematical tables, National Bureau of Standards, Washington, D. C., 1964.
2. T. Bang, Une inégalité de Kolmogorof et les fonctions presque-périodiques, Danske Vid. Selsk. Math. Fys. Medd., 19 (1941) No. 4, 28 pages.
3. W. A. Coppel, Stability and asymptotic behavior of differential equations, Heath, Boston, 1965.
4. A. Kolmogorov, On inequalities between the upper bounds of the successive derivatives of an arbitrary function on an infinite interval, Amer. Math. Soc. Translations, Series 1, 2 (1962) 233-243. This paper appeared originally in Russian in 1939.
5. E. Landau, Einige Ungleichungen für zweimal differenzierbare Funktionen, Proc. London Math. Soc., (2) 13 (1913) 43-49.
6. ———, Die Ungleichungen für zweimal differenzierbare Funktionen, Danske Vid. Selsk. Math. Fys. Medd., 6 (1925) No. 10, 49 pages.
7. ———, Über einen Satz von Herrn Esclangon, Math. Ann., 102 (1929) 177-188.
8. A. P. Matorin, On inequalities between the maxima of the absolute values of a function and its derivatives on a half-line, Amer. Math. Soc. Translations, Series 2, 8 (1958) 13-17.
9. N. E. Nörlund, Vorlesungen über Differenzenrechnung, Springer Verlag, Berlin, 1924.
10. I. J. Schoenberg, Cardinal interpolation and spline function II. Interpolation of data of power growth, MRC T. S. R. 1104, Madison, Wisconsin, 1970. To appear in J. of Approx. Theory.
11. ———, and A. Cavaretta, Solution of Landau's problem concerning higher derivatives on the halfline, MRCT. S. R. 1050, Madison, Wisconsin, 1970. Also in Proc. of the Intern. Conf. on Constructive Function Theory, Golden Sands (Varna) May 19-25, 1970, Publ. House Bulgarian Acad. Sci., Sofia, (1972) 297-308. The MRC T. S. R. 1050 is a more accurate version of the paper.
12. ———, Cardinal interpolation and spline functions VII. The behavior of cardinal spline interpolants as their degree tends to infinity, MRC T. S. R. 1184, Madison, Wisconsin, 1971. To appear in J. d'Analyse Math. (Jerusalem).
13. ———, Cardinal interpolation and spline functions VIII. To appear as an MRC T. S. Report.
14. J. N. Subbotin, On the relation between finite differences and the corresponding derivatives, Proc. Steklov. Inst. Math., 78 (1965) 24-42. Amer. Math. Soc. Translations, (1967), 23-42.

TYPES OF FULLY ORDERED GROUPS

D. P. MINASSIAN, Butler University, Indianapolis

Introduction. In this paper we shall introduce the subject of fully ordered groups and shall then examine some of these groups as discussed by Russian mathematicians in the last decade. (Apparently, translations of the Russian references are not in print.) We shall exhibit relationships among the groups introduced by the Russians and between them and more traditional classes of ordered groups. In Section 1 we shall give somewhat more than is needed as a basis for Section 2 since I assume that a study of fully ordered groups is new to most readers and so shall present some of the more interesting results (cf. [2]).

The following table of abbreviations is a quick reference to some notation we shall use, although complete definitions also appear in the text.

Table of Abbreviations

Nssg($x, y \dots$): the normal subsemigroup of a group G generated by the elements $x, y \dots$ of G .
O-group: a group which admits some full order.

O*-group: a group in which each partial order can be extended to some full order.

S-ext group: a group G in which every full order for each subgroup can be extended to some full order for G .

Sa-ext group: a group G in which every full order for each *abelian* subgroup can be extended to some full order for G .

Sn-ext group: a group G in which every full order for each *normal* subgroup can be extended to some full order for G .

San-ext group: a group G in which every full order for each *abelian normal* subgroup can be extended to some full order for G .

S*-ext group: a group G in which every *partial* order for each subgroup can be extended to some full order for G .

Sn*-ext group: a group G in which every *partial* order for each *normal* subgroup can be extended to some full order for G .

1. Preliminaries on ordered groups.

DEFINITIONS. A **partial order** for a group G is a relation \leq , "less than or equal to," on G with the usual properties: \leq is reflexive, transitive, antisymmetric ("a \leq b and b \leq a" imply $a = b$), and satisfies "a \leq b if and only if $xay \leq xby$ for each x and y in G ." A **full** or **linear** or **simple order** for G is a partial order for G such that each a and b in G are comparable (either $a \leq b$ or $b \leq a$). An example of a

Donald Minassian received his Univ. of Michigan Doctorate under E. F. Krause. Previous to that he taught mathematics, American history, and Latin in secondary schools and at two-year colleges, and took part in several summer institutes. Since taking his Degree he has been an Associate Professor at Butler University. He has worked on ordered groups and various topics in economics.
Editor.

fully ordered group is the group of additive integers under the familiar ordering; we shall give other examples below. An **O-group** (orderable group) is a group that admits some full order while the stronger **O*-group** is one for which each partial order can be extended to some full order (i.e., to a full order which may vary depending on the initial partial order).

We state some basic facts and begin with two theorems giving necessary and sufficient conditions that an abstract group be an **O-group** or an **O*-group**.

THEOREM 1.1 ([16] [17] [21]). *A group G with identity element e admits a full order (i.e., G is an **O-group**) if and only if, given a_1, \dots, a_n in G with each $a_i \neq e$, then for at least one choice of the signs $d_i = \pm 1$ one has*

$$e \notin \text{Nssg}(a_1^{d_1}, \dots, a_n^{d_n}),$$

where $\text{Nssg}(x, y, \dots)$ denotes the normal subsemigroup of G generated by x, y, \dots ; that is, $\text{Nssg}(x, y, \dots)$ consists of all products of conjugates of x, y , etc.

REMARKS. "Only if" is easy to show (cf. Remarks following Theorem 1.2 below). Theorem 1.1 is also a corollary to a result of Fuchs on the extension of a given partial order for a group to some full order for the group; see [2, p. 34 to p. 36 line 4]. Another set of conditions for a group G to admit a full order is given in [2, pp. 50–54]. Briefly, G is an **O-group** if and only if G admits a "solvable normal system" Σ of subgroups which satisfies certain additional properties. More specifically, Σ is a chain (under \subseteq) containing $\{e\}$ and G and is closed under unions, intersections, and conjugation by elements of G , and meets certain other requirements particularly regarding the "jumps." (A "jump" is a pair $C \subsetneq D$ of distinct elements of Σ with no element of Σ in between.) For example, in any jump $C \subsetneq D$ the subgroup D is normal in C , and C/D is isomorphic to a subgroup of the additive real numbers. (Hence C/D is abelian, whence the *solvable* normal system referred to above.) Note: if G is fully ordered, the elements of one such Σ are the *convex* subgroups of G ; see Definition above Proposition 1.4.

THEOREM 1.2 [22]. *A group G has the property that each partial order can be extended to some full order (i.e., G is an **O*-group**) if and only if*

- (i) if b and c are in $\text{Nssg}(a)$, then $\text{Nssg}(b)$ and $\text{Nssg}(c)$, intersect, and
- (ii) if $a \neq e$, then $e \notin \text{Nssg}(a)$, where, as above, $\text{Nssg}(x)$ denotes the normal subsemigroup of G generated by x .

REMARKS. Condition (ii) is easily seen to be equivalent to "if $a \neq e$, then the intersection of $\text{Nssg}(a)$ and $\text{Nssg}(a^{-1})$ is null." Further, if a group G satisfies (ii), then G is called **generalized torsion-free**. We note (cf. Theorem 1.1) that any **O-group** G is generalized torsion-free (and hence torsion-free), for an equation of form

$$(*) \quad \prod_{i=1}^n (x_i a x_i^{-1}) = e$$

would imply that a product of elements $x_i a x_i^{-1}$ each greater than e , or each less than e , equals e , which is impossible; in the abelian case a strong converse holds:

COROLLARY 1.3. *If G is a torsion-free abelian group, then each partial order for G extends to some full order for G (i.e., G is an O^* -group).*

Proof. In the abelian case (ii) of Theorem 1.2 is equivalent to “ G is torsion-free” since (*) of Remarks above reduces to $a^2 = e$. Also, (i) holds in any abelian group since if $b = a^m$ and $c = a^n$, then a^{mn} is in both $\text{Nssg}(b)$ and $\text{Nssg}(c)$. (Note: an easy direct proof of this corollary is given in [1].)

DEFINITION. The set of all elements g in a partially ordered group G satisfying $e \leq g$, where e is the group identity, is the (nonnegative) **cone** P .

In any ordered group the cone P determines the ordering since $a \leq b$ if and only if $e \equiv aa^{-1} \leq ba^{-1}$, that is, if and only if ba^{-1} is in P . Thus P itself is often called “a partial order” for G . It is easy to show [2, p. 13, Theorem 2] that a subset P of a group G is a cone for G if and only if P is a normal subsemigroup of G satisfying $P \cap P^{-1} = \{e\}$, where P^{-1} consists of all the inverses of the elements of P . Clearly, such P is the cone of a full order for G if and only if $P \cup P^{-1} = G$.

Each partial order P for a group G induces a partial order on any subgroup H : simply let the cone for H be $P \cap H$. Clearly, if P is a full order for G , then $P \cap H$ is a full order for H . Now suppose the subgroup H is normal in the fully ordered group G . Does P induce an order on the factor group G/H —that is, an order under which Hg is positive in G/H if and only if g is positive in G ? To help answer this question we give another definition.

DEFINITION. A subset S of the partially ordered group G is **convex** in G if, for each pair a and b of elements in S , the relation $a \leq g \leq b$ always implies that g is in S . (For example, under the familiar ordering of the real line the convex subsets are the ordinary intervals. In example 2 below, the pure imaginary numbers yi are a convex subgroup.)

PROPOSITION 1.4. *If the normal subgroup N of the fully ordered group G is convex in G , then G/N admits the induced partial ordering: Ng is positive if and only if g is positive in G .*

Proof. We define the cone for G/N to consist of the identity coset together with those cosets Ng where $g \notin N$ is positive in G ; (we must still prove this defines a cone). Now for such g the fact that N is convex in G rules out any relation of form $e \leq g \leq n$, where n is in N . Thus we conclude that g exceeds all n in N , and so all elements of such Ng are positive in the given order for G . Hence it does not matter how the coset Ng is represented, i.e., “once positive, always positive”. To verify that the requirements for a cone on G/N are met is now routine, and we omit the details. Note that the cone on G/N is the image, under the natural map $G \rightarrow G/N$, of the cone for G ; hence the term *induced* order on G/N .

The converse to Proposition 1.4 also holds; i.e., if N is a normal subgroup of the fully ordered group G such that G/N admits the induced order, then N is convex under the given order for G . For in this case the natural homomorphism $G \rightarrow G/N$ preserves order and hence, as can be easily shown for any order-preserving map between partially ordered sets (*sic*), the preimage of a convex subset is convex; (note: if both sets are fully ordered and the map is onto, then convexity is preserved in both directions). In particular, then, the kernel N of the homomorphism is a convex subset of G since N is the preimage of the “vacuously convex” identity subgroup.

Here is a related question: If G does *not* have an ordering to begin with, then when do given orders on N and G/N give rise to an order for G which induces the given orders on N and G/N ?

PROPOSITION 1.5. *If N is a normal subgroup of the group G , and if both N and G/N are partially ordered, then in order that G admit a partial order inducing those on N and G/N , it is necessary and sufficient that the cone for N be invariant in G (i.e., under conjugation by all elements of G). Clearly, if the orders for N and G/N are full, then so is the order for G .*

Proof. Necessity follows from the normality in G of N (given) and of the cone for G . For sufficiency, we let the reader verify that the conditions for a cone for G are satisfied by the union of the cone for N and the (strictly) positive cosets of G/N .

It is immediate that in any group admitting a full order the equation $x^n = a$ has at most one solution for each a (such groups are sometimes called **R-groups**): for if $c < d$ are two such solutions, then $c^n < d^n$, a contradiction. On the other hand, not all **R-groups** are **O-groups**; for an example see [14].

DEFINITION. A fully ordered group G is **archimedean** if the relation

$$a^n \leq b \text{ for all integers } n$$

always implies $a = e$.

The following two theorems show that all archimedean, and all continuous (see Theorem 1.7), fully ordered groups are essentially subgroups of the additive real numbers.

THEOREM 1.6. ([4, pp. 13–14] [2, p. 45]). *A fully ordered group G is archimedean if and only if there is an isomorphism, which preserves order, from G onto a subgroup of the naturally ordered additive group R of real numbers. That is, such G are subgroups of R .*

THEOREM 1.7 ([15], [2, p. 47]). *If $G \neq \{e\}$ is a continuous fully ordered group (i.e., each Dedekind section determines one and only one element), then there is an isomorphism, which preserves order, from G onto the naturally ordered additive group R of real numbers.*

The following result provides many examples of nonabelian fully ordered groups, as the corollary illustrates.

THEOREM 1.8 ([25]). *The free product of fully ordered groups admits a full order.*

COROLLARY 1.9. *All free groups admit a full order.*

Proof. A free group is the free product of infinite cyclic groups, each of which inherits the natural full order (or its negative) from the real numbers. (For another proof see, e.g., [2, pp. 47–49]. In fact, in an analog to the famous result that every group is the homomorphic image of a free group, B. H. Neumann and K. Iwasawa independently have proved that every fully ordered group is the image, under an order-preserving homomorphism, of a fully ordered free group; see, e.g., [2, p. 49, Theorem 9]. The proof there also shows that each partially ordered group is the image, under an order-preserving homomorphism, of a partially ordered free group where the kernel is fully ordered.)

The following are examples of fully ordered groups:

1. Any subgroup of the additive group of real numbers under the natural ordering.
2. Any subgroup of the additive group of complex numbers under the **lexicographic ordering**: $x + yi \leq 0$ if $x < 0$, or if $x = 0$ and $y \leq 0$. Note that extension of the lexicographic scheme fully orders the n -fold Cartesian product of fully ordered groups for each finite n . (More generally, a Cartesian product, under a well-ordered indexing set, of fully ordered groups admits a full order.) Note also that the order relation is non-archimedean (for instance, $ni < 1$ for all integers n yet $i \neq 0$).
3. A nonabelian example is the multiplicative group of real matrices of form

$$\begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix},$$

where the cone P consists of matrices where $a > 0$, or $a = 0$ and $b > 0$, or $a = b = 0$ and $c \geq 0$.

4. See *nonabelian example* in Section 2 and order G lexicographically.

REMARKS. All these groups are actually **O^* -groups** (each partial order extends to some full order), the first two in view of Corollary 1.3 and the last two by (for instance) the result in [9] that any 2-step solvable group which admits a full order is an **O^* -group**. (The fact that G in example 4 is 2-step solvable is shown in Section 2. The group in example 3 is 2-step solvable since it is an extension of the abelian group of all such matrices where $a = 0$ by the abelian group of matrices where $b = c = 0$; we omit the routine verification.) To illustrate, in example 1 the partial order on the integers under which the positive cone P consists only of the non-

negative *even* integers extends to a full order (the usual one) for the integers. In example 2, the usual order on the subgroup of real numbers is a partial order for the whole complex group which extends to the given full order. Like extensions apply in examples 3 and 4.

2. Russian work on ordered groups. The chart below is an attempt to portray visually the relationships discussed.

DEFINITIONS. A group G is a **S -ext group** if every full order for each subgroup of G extends to some full order for G . (For the distinguishing properties of such groups see Theorem 2.2.) Similarly G is a **Sa -ext** (resp. **Sn -ext**, **San -ext**) **group** if every full order for each abelian subgroup (resp. normal, abelian normal subgroup) of G extends to some full order for G . Group G is called a **S^* -ext group** if each *partial* order for any subgroup of G extends to some full order for G . (The Russian literature uses the notation **V -group**, **VA -group**, **VN -group**, **VAN -group**, **V^* -group** respectively for **S -ext group**, **Sa -ext group**, **Sn -ext group**, **San -ext group**, **S^* -ext group**.) An exhaustive list of references to these groups is [5], [8], [11], [12], [20], [23], [24], but we shall be somewhat selective in quoting the results therefrom.

If G satisfies any of the above definitions, then G admits a full order (i.e., G is an **O -group**) since the trivial order on the identity subgroup extends to a full order for G . Clearly, a **S^* -ext group** is a **S -ext group**, and a **S -ext group** is both a **Sa -ext** and a **Sn -ext group**, each of which is a **San -ext group**. Further, the designations **Sa^* -ext group** (every partial order for each abelian subgroup extends to some full order for G) and **San^* -ext group** (obvious meaning) are superfluous. For G is a **Sa^* -ext** (resp. **San^* -ext**) **group** if and only if G is a **Sa -ext** (resp. **San -ext**) **group** because any partial order for any abelian subgroup H of a torsion-free group G extends to a full order for H by Corollary 1.3.

On the other hand, it is unknown if a **Sn -ext group** is a **Sn^* -ext group**. However, we shall see that a *solvable* group G is a **Sn -ext group** if and only if G is a **Sn^* -ext group**.

The situation for abelian groups is very simple.

PROPOSITION 2.1. *For an abelian group G all of the designations **S^* -ext**, **S -ext**, **Sn^* -ext**, **Sn -ext**, **Sa -ext**, **San -ext** (thus also, **Sa^* -ext** and **San^* -ext** in view of their superfluosity, noted above), **O^*** , **O** , torsion-free, are equivalent.*

Proof. For any group (even if not abelian) the designation **S^* -ext** clearly implies all the others, while torsion-free is implied by all the others. Thus all we need show is that “abelian and **S^* -ext**” implies “torsion-free” and, conversely, “abelian and torsion-free” implies **S^* -ext**. In fact, the first “abelian” is superfluous. For if G , abelian or not, is a **S^* -ext group**, then the trivial order (only e is in the cone) extends to a full order for G ; hence G is an **O -group** and thus torsion-free. Conversely, suppose an *abelian* group G is torsion-free. Then a partial order P for a subgroup of G is a par-

tial order for G since P is normal in abelian G . Thus, P extends to a full order for G by Corollary 1.3. Hence G is an **S*-ext group**.

The situation for a nonabelian group G is not so cut and dried. A principal tool is this theorem of Kargapolov [5, Theorem, p. 17]:

THEOREM 2.2. *An arbitrary torsion-free group G has the property that each full order for any subgroup of G extends to some full order for G (i.e., G is a **S-ext group**) if and only if G has an abelian normal subgroup A such that (1.) the factor group G/A is abelian, and (2.) for arbitrary elements a in A and b in $G - A$ there are positive integers $m \neq n$ such that $b^{-1}a^mb = a^n$.*

REMARKS. Condition 2 shows that any such G is **metabelian**, by which we mean either one-step solvable, i.e., abelian, or two-step solvable. Necessity for Theorem 2.2 is essentially due to Terehov [23]. The fact that a , above, is not arbitrary in G but must lie in A , is omitted from the statement of the theorem, but not the proof, in [5]; that a is not arbitrary is trivial—set $a = b \neq e$. Also, it is clearly unnecessary to state, as in [5], “ $a \neq e$.”

Here is an outline of the proof:

Sufficiency: One may verify that G/A is torsion-free. Now let H be any subgroup of G with a full order P . Define $A_1 = H \cap A$ and $P_1 = P \cap A$. Clearly, A_1 is a normal subgroup of H , and so H/A_1 is a group. It is shown that A_1 is also *convex* in H (see the definition in section 1). Thus (see Proposition 1.4) under the natural map $H \rightarrow H/A_1$ the cone P induces an order \bar{P} on H/A_1 as follows: A_1x is in \bar{P} if and only if x is in P . Further, under the natural isomorphism $A_1h \leftrightarrow Ah$ between H/A_1 and HA/A , the cone \bar{P} gives rise to a partial order on $G/A \supseteq HA/A$ which extends to a full order \bar{Q} on G/A since G/A is a torsion-free abelian group and hence O^* by Corollary 1.3. Likewise the cone P extends to a full order Q_1 for the torsion-free abelian group A . Since Q_1 is invariant in G by (2), one may construct (see Proposition 1.5) a full order Q for G as follows: x in G belongs to Q if and only if x is in Q_1 , or x is not in A and Ax is in \bar{Q} . Finally, $P \subseteq Q$.

Necessity (we give but a brief sketch): The system $[N_a]$ of all convex subgroups of G forms a solvable normal system (cf. [2, pp. 50–54]). Each N_a is normal in G and each serving subgroup of N_{a+1}/N_a is normal in G/N_a . (A **serving subgroup** H of a group G is a subgroup of G such that, for each h in H and each natural number n , the equation $x^n = h$ can be solved in H if it can be solved in G .) This helps show G/Z is abelian, where Z is the intersection of the preimages of the centralizers Z_{a+1}/N_a of the factors N_{a+1}/N_a under the natural homomorphisms $G \rightarrow G/N_a$. Also, the group Z is proved abelian. Let A be a maximal abelian normal subgroup of G containing Z . Each serving subgroup of A is normal in G and so (2) holds.

REMARK. In [24] Terehov shows that any group G satisfying the conditions of

Theorem 2.2 can be embedded in a **S-ext group** where the abelian normal subgroup corresponding to A is a **divisible group**, i.e., contains all roots of each of its elements.

A nonabelian example (Terehov [24, bottom of p. 35]). Let G be the set of all ordered pairs (x, y) with integer x and rational number y under the operation

$$(x, y) \oplus (z, q) = (x + z, r^z y + q),$$

where r is any fixed positive rational number except 1. It is routine to check that G is a torsion-free nonabelian group, and that the subset $A \equiv \{(0, y), \text{ all rational } y\}$ is a normal subgroup of G such that A and G/A are isomorphic to the additive rational numbers and additive integers respectively. Further, if $(0, y)$ is in A and (z, q) is in $G - A$, then

$$(z, q) \oplus (0, y) \oplus (z, q)^{-1} = (0, r'y),$$

where $r' \equiv r^{-z}$ is a positive rational number $\neq 1$ by the definition of r and because $z \neq 0$. Thus $r' = n/m$, where $m \neq n$ are positive integers, and hence

$$m[(z, q) \oplus (0, y) \oplus (z, q)^{-1}] = n(0, y).$$

Thus G and A satisfy the hypotheses of Theorem 2.2, and G is a **S-ext group**.

From the definitions it is immediate that any subgroup of a **S-ext** (resp. **S*-ext**, **Sa-ext**) **group** inherits this property. (This remark and that on direct products of **S-ext groups**, below, have somewhat wider application in view of the chart.) Such "subgroup theorems" are of interest in ordered groups. For example, it was a classical unsolved problem if every subgroup of an **O*-group** is an **O*-group**; a counterexample is in [13]. (Trivially, each subgroup of a fully ordered group inherits a full order — the induced one — as noted in Section 1; i.e., any subgroup of an **O-group** is an **O-group**.)

Regarding "subgroup theorems," Terehov [23, p. 36] points out that a group is a **S-ext group** if every finitely generated subgroup is a **S-ext group** while Kokorin [8] gives the same result for **Sn-ext groups**. (These results have wider application by the chart; also, like results hold for **O-groups** and **O*-groups** as immediate consequences of Theorems 1.1 and 1.2 since the conditions listed there are local properties.)

Also important is whether a given class of ordered groups is closed under direct products. Kargapolov [6], and Kokorin [10] independently, show that the restricted direct product of **O*-groups** is an **O*-group**. (This generalizes a result of Kokorin [8] that the restricted direct product of **S-ext groups** is an **O*-group**, since any **S-ext group** is an **O*-group** in view of the chart to be discussed.) However, Kargapolov [6] shows that the class of **O*-groups** is *not* closed under the complete direct product, and I prove [20] that the direct product of **S-ext groups** need not be a **S-ext group**. The class of **O-groups** is easily seen to be closed under both restricted and complete direct products (under a well-ordered indexing set) by "lexicographic ordering"; see example 3 of Section 1.

In conclusion we prove, or refer to proofs of, the assertions in this chart for groups G , abelian or not (cf. Proposition 2.1):

nilpotent + San-ext \Leftrightarrow torsion-free abelian \Rightarrow

metabelian + $\text{S}^*\text{-ext}$ \Leftrightarrow metabelian + S-ext \Leftrightarrow

metabelian + $\text{Sn}^*\text{-ext}$ \Leftrightarrow metabelian + Sn-ext \Leftrightarrow

metabelian + Sa-ext \Leftrightarrow metabelian + San-ext \Leftrightarrow

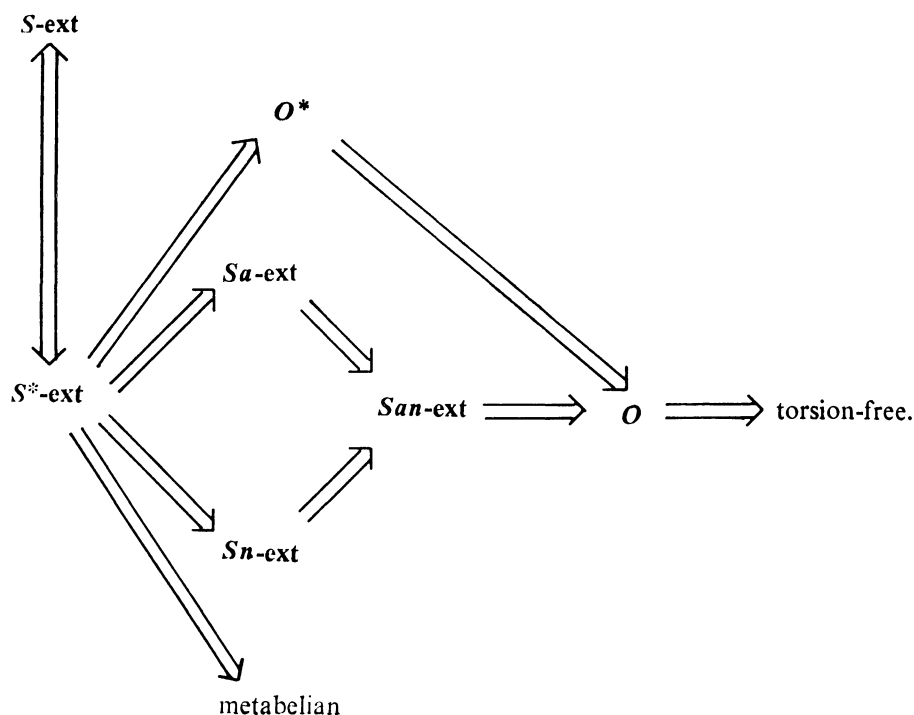


FIG. 1

“Solvable” can replace “metabelian” without loss to the validity of the chart.

Note that we have already established (or the definitions will easily establish) most of the implications on the chart. Those not yet established are the fourth and the sixth, \Leftarrow of the third, fifth and seventh, and \Rightarrow of the first, eighth and ninth.

The next theorem is like Theorem 2.2, but together the two yield important Corollary 2.4 below.

THEOREM 2.3 [11, Theorem on p. 21]. *A solvable group G is a $\text{S}^*\text{-ext}$ (resp. S-ext , Sa-ext) group if and only if G contains a torsion-free abelian normal subgroup A such that for arbitrary $g \notin A$ there are positive integers $m \neq n$ satisfying $g^{-1}a^mg = a^n$ for all a in A . (Thus the labels $\text{S}^*\text{-ext}$, S-ext and Sa-ext coincide for solvable groups.)*

COROLLARY 2.4. *The labels S -ext and S^* -ext coincide for all groups.*

Proof. Theorems 2.2 and 2.3.

Because of the importance of this last result (strangely unstated in the literature) we emphasize it: if a group G has the property that every *full* order for each subgroup of G extends to some full order for G , then every *partial* order for each subgroup of G extends to some full order for G .

REMARK. Classes O and O^* differ; all nonabelian free groups are examples of O -groups which are not O^* , but the proofs are nontrivial — see, e.g., [3]. In fact, in [6] and [7] are examples of *solvable* O -groups which are not O^* -groups.

THEOREM 2.5. (from [5] and [23]). *A solvable San -ext group is a S^* -ext group.*

REMARK. Terehov [23] actually shows, on pp. 34–35, that a solvable S -ext group meets the conditions of Theorem 2.2, but his proof applies without essential change to solvable San -ext groups; now use Theorem 2.2 and Corollary 2.4.

We now have more than enough to establish:

COROLLARY 2.6. *The labels S^* -ext, S -ext, solvable S^* -ext, solvable S -ext, solvable Sn^* -ext, solvable Sn -ext, solvable Sa -ext, and solvable San -ext coincide. ‘Metabelian’ may replace ‘solvable.’*

Every claim on the chart is now established except \Rightarrow of the first arrow. But this is Corollary 1 of Lemma 1 of [8, p. 25]. Thus any nilpotent group from the ‘new classes’ is abelian. (It is surprising that there are 2-step solvable, but not class 2 nilpotent, groups among the groups introduced in the Russian work.) Thus, nonabelian S^* -ext groups, if finitely generated, are not locally nilpotent; for example, I show in [20] that the group $\{a, b \mid ba = ab^2\}$ is a S -ext group, and hence S^* -ext by Corollary 2.4. This yields a stronger result than that of Livchak [14] that an O^* -group need not be locally nilpotent.

We have already noted that any solvable group from among the Russian groups is metabelian. This situation and that described in the previous paragraph contrast sharply from what holds for the ‘old classes’. For example, I show [19] that there are O^* -groups of arbitrary solvable length as well as unsolvable O^* -groups, while Malcev [18] shows *any* torsion-free locally nilpotent group is an O^* -group.

Finally, in a result relating the new and old groups Kopytov [12] shows that any S -ext group can be embedded in a divisible O^* -group.

This paper was completed by the author during his Faculty Fellowship at Butler University.

References

(Note: reference [2] has appeared in a revised German edition as *Teilweise geordnete algebraische Strukturen*, Studia Mathematica, Band xix, published by Vandenhoeck und Rupprecht, Göttingen, 1966. However, the footnote on p. 66, lines 5–13, is the only mention of the Russian work I have

discussed. Further, the German edition contains nothing new on the fundamentals of fully ordered groups as discussed in section 1 of this paper. Thus we refer only to the more accessible English edition.)

1. C. J. Everett, Note on a result of L. Fuchs on ordered groups, *Amer. J. Math.*, 72 (1950) 216.
2. L. Fuchs, *Partially Ordered Algebraic Systems*, Pergamon, Oxford, 1963.
3. ——— and E. Sasiada, Note on orderable groups, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.*, 7 (1964) 13–17.
4. O. Hölder, Die Axiome der Quantität und die Lehre vom Mass, *Ber. Verh. Sächs. Ges. Wiss. Leipzig. Math.-Phys. Kl.*, 53 (1901) 1–64.
5. M. I. Kargapolov, Completely ordered groups (Russian), *Algebra i Logika*, (2) 1 (1962) 16–21. MR 27 (1964) #2569.
6. ———, Fully orderable groups. I (Russian), *Algebra i Logika*, (6) 2 (1963) 5–14. MR 30 (1965) #3156.
7. ———, A. I. Kokorin and V. M. Kopytov, On the theory of orderable groups (Russian), *Algebra i Logika*, (6) 4 (1965) 21–27. MR 33 (1967) #4162.
8. A. I. Kokorin, On the theory of completely orderable groups (Russian), *Ural. Gos. Univ. Mat. Zap.*, (3) 4 (1963) 25–29. MR 32 (1966) #1271.
9. ———, On the theory of orderable groups (Russian), *Algebra i Logika*, (6) 2 (1963) 15–20. MR 30 (1965) #3157.
10. ———, Ordering a direct product of ordered groups (Russian), *Ural. Gos. Univ. Mat. Zap.*, (3) 4 (1963) 95–96. MR 29 (1965) #5938.
11. ——— and V. M. Kopytov, Certain classes of ordered groups (Russian), *Algebra i Logika*, (3) 1 (1962) 21–23. MR 27 (1964) #5840.
12. V. M. Kopytov, Completion of completely orderable groups (Russian), *Ural. Gos. Univ. Mat. Zap.*, (3) 4 (1963) 76–77. MR 32 (1966) #1272.
13. ———, On the theory of preorderable groups (Russian), *Algebra i Logika*, (6) 5 (1966) 27–31. MR 34 (1967) #4388.
14. Ja. B. Livchak, On orderable groups (Russian), *Uchen. Zap. Ural. Gos. Univ.*, 23 (1959) 11–12. MR 29 (1965) #5935.
15. F. Loonstra, Ordered groups, *Nederl. Akad. Wetensch., Proc.*, 49 (1946) 41–46.
16. P. Lorenzen, Über halbgeordnete Gruppen, *Arch. Math.*, 2 (1949) 66–70.
17. J. Los, On the existence of linear order in a group, *Bull. Acad. Polon. Sci. Cl. III*, 2 (1954) 21–23.
18. A. I. Malcev, On the completion of group order (Russian), *Trudy Mat. Inst. Steklov.*, 38 (1951) 173–175. MR 14 (1953), p. 13.
19. D. P. Minassian, On solvable O^* -groups, *Pacific J. Math.*, 39 (1971) 215–217.
20. ———, On the direct product of V -groups, *Proc. Amer. Math. Soc.*, 30 (1971) 434–436.
21. M. Ohnishi, Linear-order on a group, *Osaka Math. J.*, 4 (1952) 17–18.
22. ———, On linearization of ordered groups, *Osaka Math. J.*, 2 (1950) 161–164.
23. A. A. Terehov, Completely orderable groups (Russian), *Dokl. Akad. Nauk. SSSR*, (1) 129 (1959) 34–36. MR 22 (1961) #734.
24. ———, The structure of locally solvable completely orderable groups (Russian), *Algebra i Logika*, (2) 1 (1962) 10–15. MR 27 (1964) #2568.
25. A. A. Vinogradov, On the free product of ordered groups, (Russian), *Mat. Sb.*, 25 (1949) 163–168. MR 11 (1950), p. 157.

THE WILLIAM LOWELL PUTNAM MATHEMATICAL COMPETITION

J. H. McKAY, Oakland University

The following results of the thirty-second William Lowell Putnam Mathematical Competition held on December 4, 1971 have been determined in accordance with the regulations governing the Competition. This competition is supported by the William Lowell Putnam Intercollegiate Memorial Fund left by Mrs. Putnam in memory of her husband and is held under the auspices of the Mathematical Association of America.

The first prize, five hundred dollars, is awarded to the Department of Mathematics of **California Institute of Technology**, Pasadena, California. The members of the team were Bruce Reznick, David Smith, and Michael Yoder; to each of these a prize of one hundred dollars is awarded.

The second prize, four hundred dollars, is awarded to the Department of Mathematics of the **University of Chicago**, Chicago, Illinois. The members of the team were Robert Israel, David Saltman, and Robert Tax; to each of these a prize of seventy-five dollars is awarded.

The third prize, three hundred dollars, is awarded to the Department of Mathematics of **Harvard University**, Cambridge, Massachusetts. The members of the team were Ira Gessel, David Harbater, and Jonathan Rosenberg; to each of these a prize of fifty dollars is awarded.

The fourth prize, two hundred dollars, is awarded to the Department of Mathematics of the **University of California at Davis**, Davis, California. The members of the team were William Hamaker, Dean Hickerson, and Peter Loomis; to each of these a prize of fifty dollars is awarded.

The fifth prize, one hundred dollars, is awarded to the Department of Mathematics of the **Massachusetts Institute of Technology**, Cambridge, Massachusetts. The members of the team were Richard Arratia, David Christie, and Don Coppersmith; to each of these a prize of fifty dollars is awarded.

The six persons ranking highest in the examination, named in alphabetical order, are **Don Coppersmith**, Massachusetts Institute of Technology; **Robert Israel**, University of Chicago; **Dale Peterson**, Yale University; **Arthur Rubin**, Purdue University; **David Shucker**, Swarthmore College; **Michael Yoder**, California Institute of Technology. Each of these has been designated as a Putnam Fellow by the Mathematical Association of America and is awarded a prize of two hundred and fifty dollars.

The next four highest ranking individuals, named in alphabetical order, are *Gerald Myerson*, Harvard University, *Bruce Reznick*, California Institute of Technology; *Jonathan Rosenberg*, Harvard University, and *Angelos Tsirimokos*, Princeton University. To each of these a prize of one hundred dollars is awarded.

The following teams, named in alphabetical order, won honorable mention:

Case Western Reserve University, the members of the team were Walter Augenstein, Steven Kalikow, and Michael Somos, *University of Michigan*, the members of the team were Jonathan Glauser, Kenneth Rosen, and Dennis Stowe; *Purdue University*, the members of the team were Paul Garrett, Michael O'Donnell, and Arthur Rubin; *University of Toronto*, the members of the team were Robert Anderson, Daniel Gautreau, and Daryl Geller; *Yale University*, the members of the team were Dale Peterson, Eric Rosenthal, and Robert Weissler.

Honorable mention is given to the following thirty-two individuals, named in alphabetical order: Richard Arratia, *Massachusetts Institute of Technology*; Richard Bradley, Jr., *Massachusetts Institute of Technology*; Kenneth Brakke, *University of Nebraska at Lincoln*; Seth Breidbart, *Harvard University*; Wm. Randolph Franklin, *University of Toronto*; Paul Garrett, *Purdue University*; Daryl Geller, *University of Toronto*; John Gilbert, *University of New Mexico*; Daniel Grayson, *University of Chicago*; Charles Grinstead, *Pomona College*; Paul Hagedorn, *University of Virginia*; David Harbater, *Harvard University*; Dean Hickerson, *University of California at Davis*; David Jerison, *Harvard University*; Paul Lemke, *Rensselaer Polytechnic Institute*; Peter Loomis, *University of California at Davis*; James Lyon, *Princeton University*; Steven McKay, *Massachusetts Institute of Technology*; Peter Olver, *Brown University*; James Paulson, *Princeton University*; Richard Poppen, *Pomona College*; Arthur Rothstein, *Reed College*; Thomas Russell, *Princeton University*; David Saltman, *University of Chicago*; Eric Schechter, *University of Maryland*; Paul Selick, *University of Toronto*; David Smith, *California Institute of Technology*; Michael Somos, *Case Western Reserve University*; Dennis Stowe, *University of Michigan*; Robert Tax, *University of Chicago*; David Thornley, *University of Minnesota*; Ray White, *Princeton University*.

The other individuals who were ranked in the top one hundred, arranged by college, are: George Hardy and Daniel Kenway, *University of Alberta*; Charles Kaufman, *Bates College*; William Hamaker, *University of California at Davis*; Glenn Stevens, *University of California at Santa Barbara*; David Dummit, *California Institute of Technology*; Walter Augenstein and Steven Kalikow, *Case Western Reserve University*; Robert Hummel, Gary Miller, and David Vogan, *University of Chicago*; Joel Kleinman, *City College of New York*; David Levner and Robert Wolpert, *Cornell University*; David Kreps, *Dartmouth College*; Marcy Barge, *Fort Lewis College*; Jeffrey Dielle, William Ganong, David Garlock, Orin Gensler, Ira Gessel, Harry Porta, and Karl Strom, *Harvard University*; Jerrold Tunnell, *Harvey Mudd College*; William Van Melle, *University of Illinois*; George Cornelius, *Illinois Institute of Technology*; James Kuklinski, *LaSalle College*; Terry Andres, *University of Manitoba*; Mark Leeper, *University of Massachusetts, Amherst*; Scott Brown, David Christie, Joseph Mirzoeff, Frank Morgan, Edward Wimmers, *Massachusetts Institute of Technology*; Nozar Azarnia, *Miami University*; Jonathan Glauser, and Kenneth Rosen, *University of Michigan*; John Reiser, *Michigan State University*; Tavan Trent, *University of North Carolina*; Timothy Augustine, Steven

Garavaglia, *University of Notre Dame*; Craig Lee Huneke, *Oberlin College*; James Lawrence, *Oklahoma State University*; Bradley Jackson, *University of Oregon*; David Kallman, *University of Pennsylvania*; Eric Verheiden, *Portland State University*; Richard Enison, *Pratt Institute*; Joseph Tupper, III, *Princeton University*; Michael O'Donnell, *Purdue University*; Peter Liepa, *Queen's University, Canada*; James Alexander, *Rice University*; Jerome Eastham, Jr., *Southwestern at Memphis*; Robert Anderson, and Peter deBuda, *University of Toronto*; Thomas Templeton, *University of Wisconsin*; Eric Rosenthal and Robert Weissler, *Yale University*.¹

One thousand five hundred and sixty-nine students from three hundred and fourteen colleges and universities in the United States and Canada participated in the examination on December 4, 1971.

The Questions Committee, consisting of Warren S. Loud (chairman), Murray Klamkin, and Nathan S. Mendelsohn, prepared the problems (listed below) for the competition.

1. Students at Tel Aviv University, which is ineligible as a university outside Canada and the United States, were permitted to write the examination under the supervision of Professor Harley Flanders and have their papers graded along with the others. One of these students, Ran Donagi, would have ranked eighth in the competition and three others; Danny Berand, Joel Vodevoz, and Amnon Dalcher, would have been listed in the top hundred.

PROBLEMS. PART A

- A-1. Let there be given nine lattice points (points with integral coordinates) in three dimensional Euclidean space. Show that there is a lattice point on the interior of one of the line segments joining two of these points.
- A-2. Determine all polynomials $P(x)$ such that $P(x^2 + 1) = (P(x))^2 + 1$ and $P(0) = 0$.
- A-3. The three vertices of a triangle of sides a , b , and c are lattice points and lie on a circle of radius R . Show that $abc \geq 2R$. (Lattice points are points in the Euclidean plane with integral coordinates.)
- A-4. Show that for $0 < \varepsilon < 1$ the expression $(x + y)^n (x^2 - (2 - \varepsilon)xy + y^2)$ is a polynomial with positive coefficients for n sufficiently large and integral. For $\varepsilon = .002$ find the smallest admissible value of n .
- A-5. A game of solitaire is played as follows. After each play, according to the outcome, the player receives either a or b points (a and b are positive integers with a greater than b), and his score accumulates from play to play. It has been noticed that there are thirty-five non-attainable scores and that one of these is 58. Find a and b .
- A-6. Let c be a real number such that n^c is an integer for every positive integer n . Show that c is a non-negative integer.

PART B

- B-1. Let S be a set and let \circ be a binary operation on S satisfying the two laws

$$x \circ x = x \text{ for all } x \text{ in } S, \text{ and}$$

$$(x \circ y) \circ z = (y \circ z) \circ x \text{ for all } x, y, z \text{ in } S.$$

Show that \circ is associative and commutative.

- B-2. Let $F(x)$ be a real valued function defined for all real x except for $x = 0$ and $x = 1$ and satisfying the functional equation $F(x) + F\{(x-1)/x\} = 1 + x$. Find all functions $F(x)$ satisfying these conditions.

- B-3. Two cars travel around a track at equal and constant speeds, each completing a lap every hour. From a common starting point, the first starts at time $t = 0$ and the second at an arbitrary later time $t = T > 0$. Prove that there is a total period of exactly one hour during the motion in which the first has completed twice as many laps as the second.

- B-4. A "spherical ellipse" with foci A, B on a given sphere is defined as the set of all points P on the sphere such that $\widehat{PA} + \widehat{PB} = \text{constant}$. Here \widehat{PA} denotes the shortest distance on the sphere between P and A . Determine the entire class of real spherical ellipses which are circles.

- B-5. Show that the graphs in the x - y plane of all solutions of the system of differential equations

$$x'' + y' + 6x = 0, y'' - x' + 6y = 0 \quad (' = d/dt)$$

which satisfy $x'(0) = y'(0) = 0$ are hypocycloids, and find the radius of the fixed circle and the two possible values of the radius of the rolling circle for each such solution. (A hypocycloid is the path described by a fixed point on the circumference of a circle which rolls on the inside of a given fixed circle.)

- B-6. Let $\delta(x)$ be the greatest odd divisor of the positive integer x . Show that $\left| \sum_{n=1}^x \delta(n)/n - 2x/3 \right| < 1$, for all positive integers x .

SOLUTIONS. PART A

The number in parentheses, immediately following the problem number, is the number of participants who received a score of 8, 9 or 10 (10 is maximum possible) on the problem. In the case of A-1, A-2, B-1 and B-2, this applies to all 1569 participants. For the other problems, the count applies only to the 1039 qualifiers.

A-1 (136). The set of all lattice points can be divided into eight classes according to the parities of the coordinates, namely, (odd, odd, odd), (odd, odd, even), etc. With nine lattice points some two, say P and Q , belong to the same class. The midpoint of the segment PQ is a lattice point.

A-2 (176). $P(0) = 0$, $P(1) = [P(0)]^2 + 1 = 1$, $P(2) = [P(1)]^2 + 1 = 2$, $P(5) = [P(2)]^2 + 1 = 5$, $P(5^2 + 1) = [P(5)]^2 + 1 = 26$, etc. Thus the polynomial $P(x)$ agrees with x for more values than the degree of $P(x)$, so $P(x) \equiv x$.

A-3 (18). For a triangle with sides a, b, c , area $= A$ and circumradius $= R$ we have $abc = 4RA$. But if the vertices are lattice points the determinant formula (or Pick's Theorem or direct calculation) for the area shows that $2A$ is an integer. Hence $2A \geq 1$, so that $abc \geq 2R$. To obtain the formula $abc = 4RA$ note that if α is the angle opposite side a , then side a subtends an angle 2α at the center and $a = 2R \sin \alpha$, $A = \frac{1}{2} bc \sin \alpha$.

A-4 (49). In the expansion of $(x + y)^n(x^2 - (2 - \varepsilon)xy + y^2)$ the coefficient of $x^{k+1}y^{n+1-k}$ is

$$\begin{aligned} & \binom{n}{k-1} - (2 - \varepsilon) \binom{n}{k} + \binom{n}{k+1} \\ &= \binom{n}{k} \left\{ \frac{k}{n-k+1} + \frac{n-k}{k+1} - (2 - \varepsilon) \right\}. \end{aligned}$$

Now for fixed n consider the expression

$$\phi(k) = \frac{k}{n-k+1} + \frac{n-k}{k+1} - (2 - \varepsilon).$$

If k is taken to be a continuous positive variable

$$\phi'(k) = \frac{(n+1)\{(k+1)^2 - (n-k+1)^2\}}{(n-k+1)^2(k+1)^2}.$$

Hence $\phi'(k) = 0$ at $k = n/2$ and it follows easily that $\phi(k)$ is minimum at $k = n/2$. We needn't consider end point minima since it easily follows that for $n > 2$ the polynomial has its first two and last two coefficients positive. We may also note that if the two mid-terms in the expansion are non-positive for a given odd value of n then for the next larger value of n the mid-term remains non-positive. Hence if the mid-coefficients become positive, the first value of n for which this occurs is odd. Now if n is odd and $k = \frac{1}{2}(n+1)$ then $\phi(k) = \frac{n-1}{n+3} - 1 + \varepsilon$, and $\phi(k) > 0$ for $n > \frac{4}{\varepsilon} - 3$. If $\varepsilon = .002$, $n > 1997$ and n is odd. Hence the minimum n for which all terms are positive is 1999.

A-5 (17). The attainable scores are those non-negative integers expressible in the form $xa + yb$ with x and y non-negative integers. If a and b are not relatively prime there are infinitely many non-attainable scores. Hence $(a, b) = 1$. It will be shown that the number of non-attainable scores is $\frac{1}{2}(a-1)(b-1)$.

If m is an attainable score, the line $ax + by = m$ passes through at least one

lattice point in the closed first quadrant. Because a and b are relatively prime, the lattice points on a line $ax + by = m$ are at a horizontal distance of b . The first-quadrant segment of $ax + by = m$ has a horizontal projection of m/a and thus every score $m \geq ab$ is attainable. Every non-attainable score must satisfy $0 \leq m < ab$.

If $0 \leq m < ab$, the first-quadrant segment of the line $ax + by = m$ has a horizontal projection less than b , and so contains at most one lattice point. Thus there is a one-to-one correspondence between lattice points (x, y) with $0 \leq ax + by < ab$ in the first quadrant and attainable scores with $0 \leq m < ab$. The closed rectangle $0 \leq x \leq b$, $0 \leq y \leq a$ contains $(a + 1)(b + 1)$ lattice points, so the number of lattice points in the first quadrant with $0 \leq ax + by < ab$ is $\frac{1}{2}(a + 1)(b + 1) - 1$. This is the number of attainable scores with $0 \leq m < ab$. Hence the number of non-attainable scores in this range (which is all of them) is $ab - \frac{1}{2}(a + 1)(b + 1) + 1 = \frac{1}{2}(a - 1)(b - 1)$.

In our given example $70 = (a - 1)(b - 1) = 1(70) = 2(35) = 5(14) = 7(10)$. The conditions $a > b$, $(a, b) = 1$ yield two possibilities $a = 71$, $b = 2$ and $a = 11$, $b = 8$. Since $58 = 71(0) + 2(29)$, the first of these alternatives is eliminated. The line $11x + 8y = 58$ passes through $(6, -1)$ and $(-2, 10)$ and thus does not pass through a lattice point in the first quadrant. The unique solution is $a = 11$, $b = 8$.

A-6 (0). The case $n = 2$ shows that c is non-negative. If the ordinary mean value theorem is applied to x^c on the interval $[u, u + 1]$ there is a ξ with $u < \xi < u + 1$ such that $c \xi^{c-1} = (u + 1)^c - u^c$. For any positive integer u the right hand side is a positive integer. Now, in the case $0 < c < 1$, u could be taken large enough so $u^{c-1} < 1/c$ and so $c \xi^{c-1} < 1$. Thus the mean value theorem for the first derivative eliminates all c with $0 < c < 1$.

There is an extension of the mean value theorem which states that if $f(x)$ is k -times differentiable in $[a, b]$ then there is a ξ , $a < \xi < b$, such that $h^k f^{(k)}(\xi) = \Delta^k f(a)$, where $h = \frac{b-a}{k}$ and Δ^k is the k -th difference for intervals spaced h apart. Take k as the unique integer such that $k - 1 \leq c < k$ and apply this extension of the mean value theorem on the interval $[u, u + k]$. There is a ξ with $u < \xi < u + k$ such that

$$c(c-1)(c-2)\cdots(c-k+1)\xi^{c-k} = \Delta^k f(u).$$

The right hand side is an integer, and by taking u sufficiently large ξ^{c-1} becomes sufficiently small so that the left hand side, though non-negative, is less than 1. Hence $c(c-1)(c-2)\cdots(c-k+1) = 0$ and so $c = k - 1$.

SOLUTIONS. PART B

B-1 (735). Using the given laws we have

$$\begin{aligned} x \circ y &= (x \circ y) \circ (x \circ y) = [(x \circ y) \circ x] \circ y = [(y \circ x) \circ x] \circ y \\ &= [(x \circ x) \circ y] \circ y = (x \circ y) \circ y = (y \circ y) \circ x = y \circ x. \end{aligned}$$

From this commutative law we obtain

$$(x \circ y) \circ z = (y \circ z) \circ x = x \circ (y \circ z).$$

B-2 (314). In the given functional equation

$$(1) \quad F(x) + F\left(\frac{x-1}{x}\right) = 1 + x$$

we substitute $\frac{x-1}{x}$ for x , obtaining

$$(2) \quad F\left(\frac{x-1}{x}\right) + F\left(\frac{-1}{x-1}\right) = \frac{2x-1}{x}.$$

Also in (1), we substitute $\frac{-1}{x-1}$ for x and obtain

$$(3) \quad F\left(\frac{-1}{x-1}\right) + F(x) = \frac{x-2}{x-1}.$$

Adding (1) and (3) and subtracting (2) gives

$$(4) \quad \begin{aligned} 2F(x) &= 1 + x + \frac{x-2}{x-1} - \frac{2x-1}{x} = \frac{x^3 - x^2 - 1}{x(x-1)}. \\ F(x) &= \frac{x^3 - x^2 - 1}{2x(x-1)}. \end{aligned}$$

That $F(x)$, defined in (4), does satisfy the given functional equation is easily verified. Therefore (4) is the only solution of the problem.

B-3 (155). At time t , car 1 has completed $[t]$ laps and car 2 has completed $[t - T]$ laps. The problem is to find values of $t \geq T$ for which $[t] = 2[t - T]$.

Let $T = k + \delta$, where $0 \leq \delta < 1$, k an integer. Consider any integral interval $[m, m + 1]$ and let $m \leq t < m + 1$. Then $t = m + \varepsilon$, where $0 \leq \varepsilon < 1$. Then the equation to be solved becomes

$$[t] = m = 2[t - T] = 2[m + \varepsilon - (k + \delta)] = 2[m - k + \varepsilon - \delta].$$

Thus $m = 2(m - k)$, if $\varepsilon \geq \delta$ and $m = 2(m - k - 1)$, if $\varepsilon < \delta$. If $1 > \varepsilon \geq \delta$, then $m = 2k$ and the equation is satisfied during $[2k + \delta, 2k + 1]$, which has length $1 - \delta$.

If $0 \leq \varepsilon < \delta$, then $m = 2k + 2$ and the equation is satisfied during $[2k + 2, 2k + 2 + \delta]$ which has length δ . Therefore the total length is $1 - \delta + \delta = 1$.

Comment: The problem should have been more explicit by stating "after the start of the second car" instead of "during the motion". The solution is given for

this interpretation, whereas, if $t < T$, $[t - T]$ is negative but the second car would have completed zero laps.

B-4 (10). We take the radius of the sphere as unity and denote the constant sum $\widehat{PA} + \widehat{PB}$ by $2a$. To avoid trivial and degenerate cases we assume that $0 < \widehat{AB} < \pi$ and that $\widehat{AB} < 2a < 2\pi - \widehat{AB}$.

The case $2a > \pi$ can be reduced to the case $2a < \pi$. For, if A' and B' are the points diametrically opposite to A and B then $\widehat{PA} + \widehat{PB} = 2a$ if and only if $\widehat{PA'} + \widehat{PB'} = 2\pi - 2a$; that is, the spherical ellipses $\widehat{PA} + \widehat{PB} = 2a$ and $\widehat{PA'} + \widehat{PB'} = 2\pi - 2a$ are identical. Since $\min(2a, 2\pi - 2a) \leq \pi$, we may assume without loss of generality that $2a \leq \pi$.

Let A and B lie on the equator. There are two points V_1 and V_2 (the "vertices") on the equator which lie on the spherical ellipse. Obviously, $\widehat{V_1V_2} = 2a$. The "center" of the spherical ellipse (common midpoint of the arcs \widehat{AB} and $\widehat{V_1V_2}$) will be denoted by C .

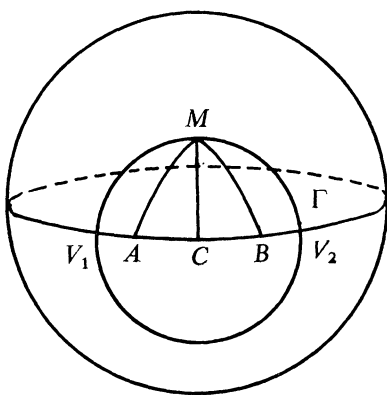


FIG. 1

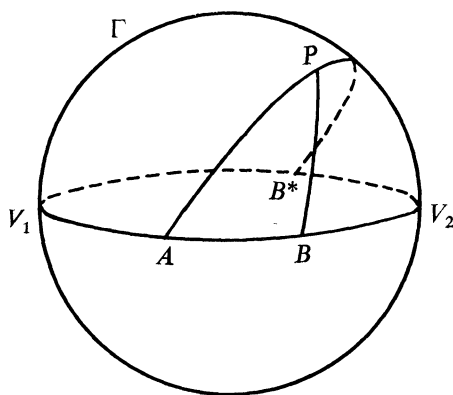


FIG. 2

We first treat the case $2a < \pi$ and show that in this case the spherical ellipse cannot be a circle. Assume it were a circle; call it Γ (see Figure 1). Γ would have to be symmetric with respect to the equatorial plane, thus lie in a plane perpendicular to the equatorial plane. Γ would also have to pass through the vertices. Therefore its spherical diameter would be $\widehat{V_1V_2} = 2a$ and its spherical radius would be equal to a . The spherical center of Γ would be C , the center of the ellipse. Let M be one of the two points on Γ which lie half-way between the two vertices. Then, since M is supposed to be a point on the spherical ellipse, $2a = \widehat{MA} + \widehat{MB} > 2\widehat{MC} = 2a$ (note that $\triangle MAC$ is a right spherical triangle with the right angle at C and with side $\widehat{MC} = a < \frac{1}{2}\pi$). Contradiction shows that the only possible spherical ellipses which are circles must occur when $2a = \pi$.

In case $2a = \pi$, V_1 and V_2 are diametrically opposite points on the equator. We shall show that the great circle Γ through the vertices and perpendicular to the equatorial plane is identical with the spherical ellipse $\widehat{PA} + \widehat{PB} = \pi$. To see this, let B^* be the reflection of B about the plane of Γ . B^* is on the equator diametrically opposite to A (see Fig. 2). Let P be an arbitrary point on the sphere, and draw the great circle through A , P and B^* . Then $\widehat{PA} + \widehat{PB^*} = \pi$. Hence, $\widehat{PA} + \widehat{PB} = \pi$ if and only if $\widehat{PB} = \widehat{PB^*}$, that is, if and only if P is on Γ . This shows that Γ is the spherical ellipse $\widehat{PA} + \widehat{PB} = \pi$, as stated above.

Thus the only circles on the sphere that are spherical ellipses are the great circles. For any given great circle Γ the foci can be any two points A and B which lie on the same great circle perpendicular to Γ , on the same side of Γ and at equal distances from Γ . The equation of any such spherical ellipse is $\widehat{PA} + \widehat{PB} = \pi$.

B-5 (7). We put $z = x + iy$. Then both differential equations can be combined into one, namely

$$(1) \quad z'' - iz' + 6z = 0.$$

This is a standard linear equation of the second order with constant coefficients and has the general solution

$$z(t) = c_1 e^{3it} + c_2 e^{-2it}.$$

The initial conditions imply $z'(0) = 0$ or $3ic_1 - 2ic_2 = 0$. We may set $c_1 = 2A$ and $c_2 = 3A$, where A is any complex number. The general solution of the given system is

$$(2) \quad z(t) = 2Ae^{3it} + 3Ae^{-2it}.$$

If $A = R e^{i\alpha}$, then a rotation of axes through the angle α produces

$$(3) \quad Z(t) = 2R e^{3it} + 3R e^{-2it}$$

or in rectangular form

$$(4) \quad \begin{aligned} X(t) &= 2R \cos(3t) + 3R \cos(2t) \\ Y(t) &= 2R \sin(3t) - 3R \sin(2t). \end{aligned}$$

This is the standard form for a hypocycloid when the radius of the rolling circle is $3R$ and the fixed circle is of radius $5R$. On time reversal it becomes the standard equations of a hypocycloid with radius of the rolling circle of $2R$ and the radius of the fixed circle of $5R$.

B-6 (52). Set

$$S(x) = \sum_{n=1}^x \frac{\delta(n)}{n}.$$

Note that $\delta(2m+1) = 2m+1$, $\delta(2m) = \delta(m)$ and that $S(2x+1) = S(2x) + 1$. Dividing the summation for $S(2x)$ into even and odd values of the index produces the following relation:

$$S(2x) = \sum_{m=1}^x \frac{\delta(2m)}{2m} + \sum_{m=1}^x \frac{\delta(2m-1)}{2m-1} = \frac{1}{2}S(x) + x.$$

If we denote $S(x) - \frac{2x}{3}$ by $F(x)$, the above relations translate into

$$F(2x) = \frac{1}{2}F(x), \text{ and } F(2x+1) = F(2x) + \frac{1}{3}.$$

Now induction can be used to show that $0 < F(x) < \frac{2}{3}$, for all positive integers x . This result is sharper than that requested.

Acknowledgements

The Director would like to acknowledge the assistance of the Questions Committee and the graders, especially Fritz Herzog, in preparing the above solutions and acknowledge the services of the following persons who were graders for the competition: J. C. Chipman, C. V. Coffman, S. E. Crick, Jr., R. A. DeVore, R. M. Dudley, W. R. Emerson, D. J. Eustice, J. Froemke, R. A. Gambill, L. J. Green, R. C. Hamelink, M. Hausner, F. Herzog, L. M. Kelly, B. B. Lieberman, W. S. Loud, D. A. Malm, E. A. Nordhaus, R. Pollack, I. Schochetman, M. E. Shanks, J. P. Williams, E. T. Wong.

MATHEMATICAL NOTES

EDITED BY ROBERT GILMER

The present backlog for this Department is substantial. Until further notice, new manuscripts cannot be accepted. This moratorium will probably continue until June 1, 1973; authors are requested to hold their manuscripts pending a further announcement.

THE AREA OF A HYPERSPHERE IN RIEMANNIAN SPACE

B. A. FUSARO, Queens College, N. C. & National Taiwan Normal University

1. Introduction. In Euclidean m -space the surface area Ω_m of a (hyper)sphere with radius t is given by [1, p. 303]

$$(1) \quad \Omega_m = \omega_m t^{m-1}, \quad \omega_m = 2\pi^{m/2}/\Gamma(m/2),$$

where ω_m denotes the area of a unit sphere in E^m . For example, from the value $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ we get the familiar $\omega_2 = 2\pi$ and $\omega_3 = 4\pi$.

We shall consider a Riemannian m -space and ask: Is the area Ω_m of a sphere

of fixed radius independent of the location of the sphere in the space? Just as a physicist employs the concept of a test electric charge, or magnetic pole to probe physical space, so a mathematician can arm himself with a test sphere to examine geometrical space. More precisely, we associate with each point P of m -space a closed spherical neighborhood of radius t and (in order to avoid self-intersecting or "incomplete" neighborhoods) consider only those that are homeomorphic to a closed m -sphere in E^m . We then ask: Is $\Omega_m = \Omega_m(P, t)$ independent of P for every fixed t ?

In E^m the answer is evidently *yes*. Another space that is of interest is the space K^m of constant curvature c . This space, which is isometric to E^m when $c = 0$, has a somewhat involved definition via a curvature tensor. However, for $c > 0$ we can interpret K^m as the surface of an $(m + 1)$ -sphere of radius $1/\sqrt{c}$ in E^{m+1} (see [3]). It follows from the geometry that the area of a sphere is independent of its location in a space of positive constant curvature. In fact, the independence property holds for all c with

$$(2) \quad \Omega_m = \omega_m \begin{cases} [\sin(t\sqrt{c})/\sqrt{c}]^{m-1} & 0 < c < (\pi/t)^2 \\ [\sinh(t\sqrt{-c})/\sqrt{-c}]^{m-1} & c < 0. \end{cases}$$

Fulton [3] has recently given a simplified derivation of (2).

The Harmonic spaces H^m of Copson and Ruse (1939) enter the scene quite naturally as the next superspace. These spaces specialize to K^m for $m = 2, 3$. [See 5]

It will be shown in this note that it follows readily from known results in H^m theory, the use of Riemannian normal coordinates, and a simplified definition (3) of Ω_m that spheres in H^m do have this independence property. The converse property is discussed in the last section.

2. Riemannian space and Riemannian normal coordinate systems (RNCS). We begin with an m -dimensional differentiable manifold of class C^n . Briefly this is a Hausdorff $(T-2)$, locally Euclidean, connected space that, via suitable homeomorphic maps into E^m , allows at each point the erection of a coordinate system (CS) and C^n transformations to other CS [4 or 6]. We introduce on the manifold a CS and convert it to a metric space via the symmetric, positive-definite, C^2 quadratic form

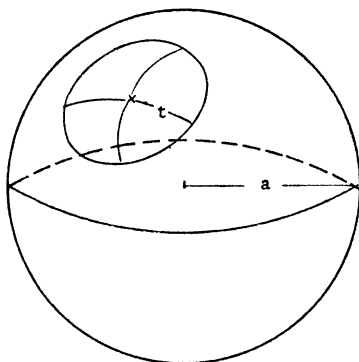
$$ds^2 = g_{ij}(x)dx^i dx^j, \quad x = (x^1, x^2, \dots, x^m).$$

The convention of summing from 1 to m on repeated indices is used throughout this note. The CS and metric will be denoted by

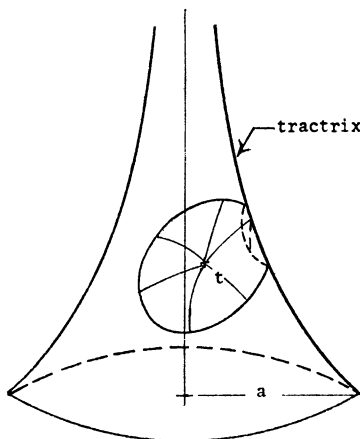
$$(x; \sqrt{g}), \quad g = \det(g_{ij}).$$

A transformation from $(x; \sqrt{g})$ to $(y; \sqrt{h})$ is given by

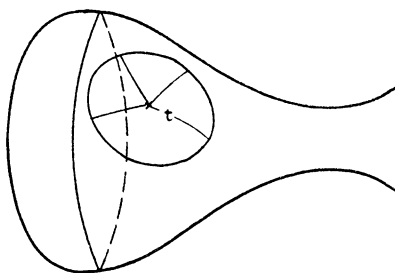
$$y^i = f^i(x), \quad h_{ij} = (\partial x^p / \partial y^i) (\partial x^q / \partial y^j) g_{pq}.$$



Positive curvature $1/a^2$
 $\Omega_2 = 2\pi \cdot a \sin(t/a)$



Negative curvature $-1/a^2$
 $\Omega_2 = 2\pi \cdot a \sinh(t/a)$



Mixed curvature
 $\Omega_2 = 2\pi \cdot \sqrt{g(t)}$

The metric is given by $r = |Q - P| = \min \int_P^Q (ds/d\tau) d\tau$, where the minimum is taken over all continuously differentiable parametrized arcs $x = x(\tau)$ connecting P to a neighboring point Q . A minimizing arc is a geodesic. A (geodesic hyper) sphere in this m -space can then be defined as in Euclidean space.

A coordinate system can be chosen so that $ds^2 = \delta_{ij} dx^i dx^j$ if and only if the space is E^m . However, it is always possible on a class C^3 differentiable manifold to choose a system $(y; \sqrt{h})$ so that $h_{ij} = \delta_{ij}$ at the origin and so that equations of geodesics issuing from the origin have the linear form $y^i = \beta^i r$, with $|\beta| = 1$, and where r denotes geodesic distance measured from the origin. This system is known as a **Riemannian normal coordinate system**, abbreviated **RNCS**. A property of a RNCS is that $h_{ij} y^i = \delta_{ij} y^j$ along a geodesic so that the square of the distance from the origin measured along a geodesic takes the form $r^2 = \delta_{ij} y^i y^j$. See [4, pp. 149, 307] for a full discussion. These special coordinate systems are very useful

because they put many of the results of Riemannian geometry in familiar Euclidean form.

3. The surface Ω_m of a sphere in Riemannian space. Consider, in a Riemannian m -space, a sphere with center P and radius t , and let Q denote a variable point of the sphere. In the system $(x; \sqrt{g})$ the volume of this sphere is

$$\int_{|x-\xi| < t}^{(m)} \sqrt{g} dx \quad P = P(\xi), \quad Q = Q(x).$$

The area of its surface will be defined by the equation

$$(3) \quad \Omega_m = \Omega_m(P, t) = d/dt \int_{|x-\xi| < t}^{(m)} \sqrt{g} dx.$$

If the space is Euclidean then (3) reduces to (1). This definition by-passes the difficulties attending the parametrization of a surface.

We shall assume that $(x; \sqrt{g})$ is referred to a RNCS with origin at P and choose $\xi = 0$. The indicated differentiation in (3) can be explicitly carried out after a transformation to a geodesic polar system $(r, \theta; \sqrt{\gamma})$. If we interpret our RNCS as rectangular coordinates, this transformation takes the form of the usual one for polar coordinates [2, p. 65]. The typical case $m = 4$, written in subscript notation, is

$$x_1 = r \sin \theta_1 \sin \theta_2 \cos \theta_3$$

$$x_2 = r \sin \theta_1 \sin \theta_2 \sin \theta_3$$

$$x_3 = r \sin \theta_1 \cos \theta_2$$

$$x_4 = r \cos \theta_1$$

with θ_i in $[0, \pi]$ for $i = 1, 2$ and θ_3 in $[0, 2\pi)$. Here r denotes the Euclidean distance from the origin P and also represents geodesic distance $|Q - P|$ in the original RNCS. The angles θ_k are measured at the origin and have their usual Euclidean meaning.

The Jacobian J of the above transformation has the form $J = r^m \Phi(\theta)$, and satisfies the relation $\sqrt{\gamma} = |J| \sqrt{g}$. The volume integral over the sphere $|x| < t$ can now be written as

$$\int^{(m)} \sqrt{g} dx = \int^{(m)} \sqrt{\gamma} dr d\theta = \int^{(m)} \sqrt{g} |J| dr d\theta = \int_0^t r^{m-1} dr \oint^{(m-1)} \sqrt{g} \Phi(\theta) d\theta.$$

A Harmonic space H^m can be characterized by the property that g is radially symmetric [5, p. 35]. That is, the coordinate variable x and the parameter ξ enter $g = \det(g_{ij})$ only via $r = |P - Q|$, so that one g -function serves for the whole space. From definition (3) we have

$$\Omega_m = d/dt \int_0^t \sqrt{g} r^{m-1} dr \oint^{(m-1)} \Phi(\theta) d\theta = t^{m-1} \sqrt{g(t)} \oint^{(m-1)} \Phi(\theta) d\theta.$$

The angle integral in this expression for Ω_m is the ordinary E^m unit surface element ω_m in (1), as is seen by letting $H^m = E^m$ and choosing $t = 1$. Therefore,

$$(4) \quad \Omega_m = \sqrt{g(t)} t^{m-1} \omega_m \text{ in } H^m \text{ (referred to a RNCS),}$$

so that in a *Harmonic space* the area of a sphere of given radius is independent of its location in the space. If the space is of constant curvature, then

$$\sqrt{g(t)} = [\sin(t\sqrt{c})/(t\sqrt{c})]^{m-1}$$

and the above expression for Ω_m reduces to (2), as it should [5, p. 30].

Can the above argument be reversed to show that if a test sphere has the independence property then g is radially symmetric so that the space is Harmonic:

$$\Omega_m(P, t) = \Omega_m(t) \Rightarrow g(r, \theta) = g(r)?$$

No, because it can happen that the angle θ enters g at each point P in such a way as to be integrated away.

4. The mean-value as an even function of the radius. Let $M = M(t, \xi; f)$ denote the mean-value of a continuous function f averaged over a sphere with radius t and center $P(\xi)$

$$\Omega_m \cdot M = \oint^{(m-1)} f(x) dS_x, \quad |x - \xi| = t.$$

After a transformation to a RNCS with origin at P , this expression takes the form

$$\Omega_m \cdot M = \oint f(\xi + \alpha t) dS_x, \quad x = \xi + \alpha t, \quad |\alpha| = 1,$$

which is defined for negative t . Now assume the space is Harmonic and apply equation (4) to get

$$\omega_m \cdot M = \oint f(\xi + \alpha t) d\omega_\alpha \quad |\alpha| = 1.$$

The usual Euclidean argument then yields that M is an even function of t .

5. A tempting conjecture. If a test sphere indicates that a space has the independence property, is that space Harmonic? The tempting affirmative conjecture, if correct, yields a simple geometric characterization of H^m . This converse question does not appear to be an easy one to answer. It is worth knowing whether the answer is *yes* even for a space K^m of constant curvature, especially because $H^m = K^m$ for the cases $m = 2, 3$, as was remarked earlier. For the case $m = 2$, at least, the answer is *yes*. If Ω_2 is independent of the center P of the sphere, it follows from

the Puiseux-Bertrand formula [4, p. 151] that for curvature $c(P)$

$$\pi \cdot c(P) = 3 \lim_{t \rightarrow 0} (2\pi t - \Omega_2)/t^3$$

that the space is of constant curvature.

Supported in part by NSF grant GP 1834.

References

1. R. Courant, *Differential and Integral Calculus*, vol. II, Interscience, New York, 1964.
2. L. E. Blumenson, A derivation of n -dimensional spherical coordinates, this MONTHLY, 67 (1960) 63–66.
3. C. M. Fulton, Hyperspheres in spaces of constant curvature, this MONTHLY, 76 (1969) 43–44.
4. E. Kreyszig, *Differential Geometry and Riemannian Geometry*, University Press, Toronto, 1968.
5. H. S. Ruse, A. G. Walker, T. J. Willmore, *Harmonic Spaces*, Edizioni Cremonese, Rome, 1961.
6. T. J. Willmore, *An Introduction to Differential Geometry*, Oxford University Press, London, 1959.

AN AREA THEOREM FOR SCHLICHT FUNCTIONS

J. L. ULLMAN, University of Michigan

The theorem proved in this paper was observed by Prof. H. Alexander, of the University of Michigan, to be a consequence of theorems concerning functions of several complex variables found in Rutishauser [2, p. 257, p. 259]. The content of the theorem states a simple and, we believe, interesting property concerning univalent conformal maps, and we feel that a proof based on tools on one-variable complex analysis will be of interest.

THEOREM. *Let $w = f(z)$ be analytic in the domain $\Sigma_z = \{z : |z| < 1\}$, univalent, and let $f(0) = 0$. Let $\Sigma_w = \{w : |w| < 1\}$, let $S = f(\Sigma_z) \cap \Sigma_w$, and let $A = f^{-1}(S)$. It is then true that (a)*

$$(1) \quad \text{Area}(A) + \text{Area}(S) \geq \pi,$$

and (b) that no larger constant can be used on the right side of (1).

Proof. An example shows that (b) holds. Namely, let $w = \lambda z$, $|\lambda| < 1$. Then we have $S = \{w : |w| < |\lambda|\}$, $A = \Sigma_z$, and $\text{Area}(A) + \text{Area}(S) = (1 + |\lambda|^2)\pi$. Thus (b) is established since $|\lambda|$ can be made arbitrarily small.

The proof of (a) is divided into two cases. In Case I, we assume that $f(z)$ is analytic and univalent on the set $\bar{\Sigma}_z = \{z : |z| \leq 1\}$ and in Case II, the general case is considered.

Case I. If $\sigma_z = \{z : |z| = 1\}$, then $f(\sigma_z)$ is an analytic Jordan curve, and either intersects $\sigma_w = \{w : |w| = 1\}$ a finite number of times or coincides with σ_w . In the latter case, $f(z) = \alpha z$, $|\alpha| = 1$, and the theorem is true, so we consider the first situation. The set S need not be a connected set, but S_0 , the component of S containing $w = 0$, is simply connected and bounded by a simple piecewise analytic curve, $\Gamma(S_0)$. We only use the fact that $\Gamma(S_0)$ is a Jordan curve. Thus $A_0 = f^{-1}(S_0)$ is a simply connected subset of Σ_z containing $z = 0$, and is bounded by a Jordan curve $\Gamma(A_0)$. Since $S_0 \subset S$ and $A_0 \subset A$, it is sufficient to prove

$$(2) \quad \text{Area}(A_0) + \text{Area}(S_0) \geq \pi.$$

Because of the stated properties of A_0 , we know by the Riemann mapping theorem and Caratheodory's theorem on the boundary behavior of the mapping function (Hille [1, p. 320, p. 360]), that there is a function $h(t)$ analytic in $\Sigma_t = \{t : |t| < 1\}$ and continuous in $\bar{\Sigma}_t = \{t : |t| \leq 1\}$ such that Σ_t is mapped univalently onto A_0 and $\sigma_t = \{t : |t| = 1\}$ is mapped univalently onto $\Gamma(A_0)$. Furthermore, the function $g(t) = f(h(t))$ is analytic in Σ_t and continuous in $\bar{\Sigma}_t$ and maps Σ_t univalently onto S_0 and σ_t univalently onto $\Gamma(S_0)$. Thus if $h(t) = \sum_{n=1}^{\infty} a_n t^n$ and $g(t) = \sum_{n=1}^{\infty} b_n t^n$, we have (Hille [1, p. 360])

$$(3) \quad \text{Area}(A_0) + \text{Area}(S_0) = \pi \sum_{n=1}^{\infty} n(|a_n|^2 + |b_n|^2).$$

Since $h(t)$ and $g(t)$ are continuous on σ_t , we also have the relations (Hille [1, p. 360])

$$(4) \quad \frac{1}{2\pi} \int_0^{2\pi} |h(e^{i\theta})|^2 d\theta = \sum_{n=1}^{\infty} |a_n|^2, \quad \frac{1}{2\pi} \int_0^{2\pi} |g(e^{i\theta})|^2 d\theta = \sum_{n=1}^{\infty} |b_n|^2.$$

If $e_1 = \{e^{i\theta} : |h(e^{i\theta})| = 1\}$, $e_2 = \{e^{i\theta} : |g(e^{i\theta})| = 1\}$, we find that

$$(5) \quad \begin{aligned} \frac{1}{2\pi} \int_0^{2\pi} |h(e^{i\theta})|^2 d\theta &\geq \frac{\text{meas}(e_1)}{2\pi} \\ \frac{1}{2\pi} \int_0^{2\pi} |g(e^{i\theta})|^2 d\theta &\geq \frac{\text{meas}(e_2)}{2\pi}, \end{aligned}$$

where $\text{meas}(e_j)$ indicates the linear Lebesgue measure of

$$\{\theta : e^{i\theta} \in e_j, 0 \leq \theta \leq 2\pi\}, \quad j = 1, 2.$$

We use the fact in (5) that $|h(e^{i\theta})| \geq 0$ and $|g(e^{i\theta})| \geq 0$. Once we show

$$(6) \quad \text{meas}(e_1) + \text{meas}(e_2) \geq 2\pi,$$

the combination of (3), (4), (5), and (6) yields (2), so we proceed to the proof of (6).

If $\text{meas}(e_1) = 2\pi$, (6) follows, so we consider the case that this is not so. There

will then be a point on σ_r not in e_1 . If we can show that such a point must be in e_2 , (6) will hold. Assume then that $e^{i\theta_1} \notin e_1$, so that $|h(e^{i\theta_1})| < 1$. Now $g(e^{i\theta_1})$ is a point of $\Gamma(S_0)$. The set $\Gamma(S_0)$ consists of arcs of σ_w and arcs of $f(\sigma_z)$. Assume next that $e^{i\theta_1} \notin e_2$. This means that $|g(e^{i\theta_1})| < 1$. Thus $g(e^{i\theta_1})$ is a point of $f(\sigma_z)$. On the other hand, $g(e^{i\theta_1}) = f(h(e^{i\theta_1}))$, and since $|h(e^{i\theta_1})| < 1$, $f(h(e^{i\theta_1}))$ is an interior point of $f(\Sigma_z)$ and hence cannot be a point of $f(\sigma_z)$. This contradiction completes the proof for Case I.

Case II. If $w = f(z)$ is analytic in the domain Σ_z , univalent and satisfies $f(0) = 0$, then $f_r(z) = f(rz)$ satisfies the requirements of Case I when $0 < r < 1$. Thus if $S_r = f_r(\Sigma_z) \cap \Sigma_w$ and $A_r = f_r^{-1}(S_r)$, we have by (1)

$$(7) \quad \text{Area}(A_r) + \text{Area}(S_r) \geq \pi.$$

It remains to show

$$(a) \lim_{r \uparrow 1} \text{Area}(S_r) = \text{Area}(S) \quad \text{and} \quad (b) \lim_{r \uparrow 1} \text{Area}(A_r) = \text{Area}(A)$$

and the proof is complete. Now S_r is an open set, increasing as $r \uparrow 1$ and exhausts S , thus establishing (a). In addition

$$f_r^{-1}(S_r) = \frac{1}{r} f^{-1}(S_r),$$

and so (b) is established by letting r tend to one, and the proof of Case II is complete.

The methods of this paper have been enlarged upon, and have led to the following theorem which will appear in [3].

THEOREM. *If $f(z)$ is continuous in $\overline{\Sigma_z}$, analytic in Σ_z and satisfies $f(0) = 0$, then $\frac{1}{2} \int_0^{2\pi} |f(e^{i\theta})|^2 d\theta \leq A(D)$, where $D = f(\Sigma_z)$ and multiplicity is not counted in measuring area.*

References

1. E. Hille, *Analytic Function Theory*, Vol. II, Ginn and Company, New York, 1962.
2. H. Rutishauser, Über Folgen und Scharen von analytischen und meromorphen Functionen mehrerer Variablen, sowie von analytischen Abbildungen, *Acta Mathematica*, 83 (1950) 249–325.
3. H. Alexander, B. A. Taylor and J. L. Ullman, Areas of Projections of Analytic Sets, *Inventiones Mathematicae*, Fasc. 4., 16 (1972) 335–341.

ON SET POINTS OF DISCONTINUITY

O. T. ALAS, University of São Paulo

In this note we shall prove a theorem which is a generalization of a well-known theorem on metric spaces [2].

THEOREM. *Let X be a topological space, (Y, \mathcal{U}) be a Hausdorff uniform space of weight m , (f_n) be a sequence of continuous functions of X into Y and f be a function of X into Y such that $f(x) = \lim f_n(x)$ for every $x \in X$. Then the set of*

points of discontinuity of f is the union of at most m (\aleph_0) nowhere dense subsets of X if m is infinite (respectively, if m is finite).

Before turning to the proof, let us recall some definitions. The weight of a uniform space (Y, \mathcal{U}) is the least cardinal number such that the uniformity \mathcal{U} has a basis of this cardinality. We may assume (see [1] page 186) that the elements of this basis are symmetric, closed subsets of the product space $Y \times Y$. A subset of X is nowhere dense if the interior of its closure is empty.

Proof of the theorem. Our proof follows that which appears in [2].

Let us denote by B a basis of the uniformity \mathcal{U} , whose cardinality is m and whose elements are symmetric closed subsets of $Y \times Y$.

For each $U \in B$, let us denote by $D(U)$ the set of the points $x \in X$ such that $f(V) \times f(V) - U \neq \emptyset$ for every neighborhood V of x . Thus, denoting by D the set of the points of discontinuity of f , we have that

$$D = \bigcup \{D(U) \mid U \in B\}.$$

We shall prove that each set $D(U)$ is the union of a countable number of nowhere dense subsets of X .

Fix $U \in B$ and $W \in B$ such that $W \circ W \circ W \subset U$. For each natural number $k \geq 1$ put

$$A_k = \{x \in X \mid (f_n(x), f_k(x)) \in W, \forall n \geq k\}.$$

The identity $X = \bigcup \{A_k \mid k = 1, 2, \dots\}$ holds by virtue of the convergence of the sequence (f_n) . Each set A_k is closed because the f_n are continuous, and W is closed in $Y \times Y$. Thus $D(U) = \bigcup \{D(U) \cap A_k \mid k = 1, 2, \dots\}$ and every set $D(U) \cap A_k$ is nowhere dense. Indeed, if x belongs to the interior of A_k , since f_k is continuous, there is an open neighborhood G of x , contained in A_k , such that $f_k(G) \times f_k(G) \subset W$. If $y, z \in G$, then $(f(y), f_k(y)) \in W$, $(f(z), f_k(z)) \in W$, and $(f_k(y), f_k(z)) \in W$. So $(f(y), f(z)) \in U$ and x does not belong to $D(U)$. The proof is completed.

References

1. N. Bourbaki, *Topologie Générale*, livre 3, chapitres 1 & 2, Hermann, Paris, 1965.
2. K. Kuratowski, Sur les fonctions représentables analytiquement et les ensembles de première catégorie, *Fund. Math.*, 5 (1924) 75–86.

GENERALIZED FIBONACCI NUMBER TRIPLES

A. G. SHANNON, New South Wales Institute of Technology, Sydney, Australia, and

A. F. HORADAM, University of New England, Armidale, Australia and University of Reading, England

1. Introduction. It is possible to relate the results of Teigen and Hadwin [5] to the generalized sequence of numbers, $\{w_n\}$, investigated by Horadam [3], and, at

the same time, to generalize the Fibonacci number triples related to $\{H_n\}$ previously studied by Horadam in this MONTHLY [1], [2].

$\{w_n\}$ satisfies the general second order recurrence relation

$$(1.1) \quad w_n = pw_{n-1} - qw_{n-2} \quad (n \geq 2)$$

with general initial conditions $w_0 = a$, $w_1 = b$, and where p and q are arbitrary integers. When $p = -q = 1$, $\{w_n\} \equiv \{H_n\}$.

2. Lemmas.

$$(2.1) \quad (p^2 - q)w_{n+2} - pw_{n+3} = q^2w_n.$$

$$(2.2) \quad (p^2 - q)w_{n+2} + pw_{n+3} = 2(p^2 - q)w_{n+2} - q^2w_n.$$

$$\begin{aligned} \text{Proof of (2.1).} \quad & (p^2 - q)w_{n+2} - pw_{n+3} \\ &= (p^2 - q)w_{n+2} - p^2w_{n+2} + pqw_{n+1} \quad \text{by (1.1)} \\ &= -pqw_{n+1} + q^2w_n + pqw_{n+1} \quad \text{by (1.1)} \end{aligned}$$

which gives the required result.

The proof of (2.2) follows immediately from (2.1).

3. Theorems.

$$(3.1) \quad \{(p/q^2)w_n w_{n+3}\}^2 + \{2Pw_{n+2}(Pw_{n+2} - w_n)\}^2 = \{w_n^2 + 2Pw_{n+2}(Pw_{n+2} - w_n)\}^2$$

where $P = (p^2 - q)/2q^2$.

The three numbers in the Pythagorean-type formula (3.1) are called a *generalized Fibonacci triple*.

(3.2) *All Pythagorean triples are generalized Fibonacci triples.*

Proof of (3.1). Multiply the corresponding sides of (2.1) and (2.2):

$$(p^2 - q)^2 w_{n+2}^2 - p^2 w_{n+3}^2 = 2(p^2 - q) q^2 w_n w_{n+2} - q^4 w_n^2.$$

Divide through by q^4 and rearrange to obtain

$$\{(p/q^2)w_{n+3}\}^2 = w_n^2 + 4Pw_{n+2}(Pw_{n+2} - w_n).$$

Multiply through by w_n^2 and add $4P^2w_{n+2}^2(Pw_{n+2} - w_n)^2$ to each side:

$$\begin{aligned} & \{(p/q^2) \cdot w_n w_{n+3}\}^2 + \{2Pw_{n+2}(Pw_{n+2} - w_n)\}^2 \\ &= (w_n^2)^2 + 2\{2Pw_{n+2}(Pw_{n+2} - w_n)\} (w_n^2) + \{2Pw_{n+2}(Pw_{n+2} - w_n)\}^2 \\ &= \{w_n^2 + 2Pw_{n+2}(Pw_{n+2} - w_n)\}^2, \text{ as required.} \end{aligned}$$

Proof of (3.2). Put $a = t(x - y)$, $b = t((1 + Pq)x - Pqy)/Pp$ in $\{w_n\}$ to obtain

the sequence

$$(3.3) \quad t(x - y), t((1 + Pq)x - Pqy)/Pp, tx/P, t(x + y)q^2/p, \dots$$

For $n = 0$, (3.1) and (3.3) give $t^4(x^2 - y^2)^2 + (2t^2xy)^2 = t^4(x^2 + y^2)^2$, which proves the theorem.

4. Examples. When $p = -q = 1$, $P = 1$, and (3.1) reduces to

$$(4.1) \quad (H_n H_{n+3})^2 + (2H_{n+1} H_{n+2})^2 = (H_n^2 + 2H_{n+1} H_{n+2})^2$$

which is equation (3) of [2]. (3.1) in fact agrees with equation (2.2) of [3], namely,

$$\begin{aligned} & [(pw_{n+1} - qw_n)^2 - w_{n+1}^2]^2 + [2w_{n+1}(pw_{n+1} - qw_n)]^2 \\ & = [(pw_{n+1} - qw_n)^2 + w_{n+1}]^2 \end{aligned}$$

but (3.1) above is in a form which can be generalized for recurrence relations of order higher than two as in Shannon and Horadam [4].

More specifically, when $t = 1$, and $p = -q = P = 1$, the formula in the proof of (3.2) leads to

$$(x^2 - y^2)^2 + (2xy)^2 = (x^2 + y^2)^2$$

which gives primitive pythagorean triples. For example, when $x = 2$, $y = 1$, we obtain the primitive triple 3, 4, 5 from the Fibonacci sequence 1, 1, 2, 3, ... (derived from (3.3)); when $x = 3$, $y = 2$, we obtain the primitive triple 5, 12, 13 from the sequence 1, 2, 3, 5, ... Of course, if $t = 3$ (say) and $p = -q = P = 1$, we obtain the nonprimitive triple 9, 12, 15 [= 3 (3, 4, 5)] from the sequence 3, 3, 6, 9, ... [= 3 (1, 1, 2, 3, ...)].

5. Methods of Teigen, Hadwin and Horadam. We conclude by showing how (4.1) and the method of Teigen and Hadwin are related. Teigen and Hadwin proved that a Pythagorean triple (a, b, c) can be represented by

$$(5.1) \quad a = x + z, b = y + z, c = x + y + z,$$

where x, y, z satisfy

$$(5.2) \quad x, y, z \text{ are positive, } 2xy = z^2, z \text{ is even.}$$

If we set $x = H_n^2$, $z = 2H_n H_{n+1}$, then $z^2 = 2H_n^2(2H_{n+1}^2)$, so that $y = 2H_{n+1}^2$ from (5.2). If we use (5.1) and (1.1) with $p = -q = 1$, we find that

$$\begin{aligned} a &= H_n^2 + 2H_n H_{n+1} = H_n H_{n+3}, \\ b &= 2H_{n+1}^2 + 2H_n H_{n+1} = 2H_{n+1} H_{n+2}, \\ c &= H_n^2 + 2H_{n+1}^2 + 2H_n H_{n+1} = H_n^2 + 2H_{n+1} H_{n+2}. \end{aligned}$$

Thus (a, b, c) is related to the Fibonacci triple of (4.1).

References

1. A. F. Horadam, A generalized Fibonacci sequence, this MONTHLY, 68 (1961), 455-459.
2. ———, Fibonacci number triples, this MONTHLY, 68 (1961), 751-753.
3. ———, Special properties of the sequence $w_n(a, b; p, q)$, Fibonacci Quart., 5 (1967) 424-434.
4. A. G. Shannon and A. F. Horadam, A generalized Pythagorean theorem, Fibonacci Quart., 9 (1971) 307-312.
5. M. G. Teigen and D. W. Hadwin, On generating Pythagorean triples, this MONTHLY, 78 (1971), 378-379.

AMBIVALENCE IN ALTERNATING SYMMETRIC GROUPS

CLAIRE PARKINSON, Burlington, Vermont

Referring to Higgins and Ballew [1, p. 274] we have the following

DEFINITION. An **ambivalent element** of a group is one which is conjugate to its inverse; an **ambivalent group** is one all of whose elements are ambivalent.

Certain groups are readily seen to be ambivalent, such as all full symmetric and all dihedral groups. Others are readily seen to be nonambivalent, such as all odd order groups of more than one element. In this article the ambivalence of alternating symmetric groups will be examined, with the result reached that the only such ambivalent groups are A_1 , A_2 , A_5 , A_6 , A_{10} , and A_{14} . Throughout, S_n denotes the full symmetric group on n letters and A_n its subgroup consisting of all even permutations.

LEMMA. S_n is ambivalent.

Proof. Immediate from Herstein [2, pp. 75-76] since all elements in S_n with the same cycle decomposition are conjugate.

The lemma establishes that each element $x \in A_n$ has a conjugating element $t \in S_n$ taking $x \rightarrow x^{-1} = t^{-1}xt$. To show the ambivalence of x in A_n we must show that some such t is in A_n as well as S_n .

THEOREM 1. The nonambivalent elements of A_n are precisely the elements x with cycle decomposition $\{I_1, I_2, \dots, I_m\}$ satisfying the following three restrictions:

1. Each I_i is odd.
2. $I_i = I_j \Rightarrow i = j$.
3. $\frac{1}{2}(n - m)$ is odd.

Proof. The method of proof will be to show first (a) if x is nonambivalent then 1 and 2 hold and second (b) if 1 and 2 hold then x is nonambivalent if and only if 3 holds also.

(a) If some I_i is even, let $(x_1 x_2 \dots x_{2d})$ be any cycle of even order in the disjoint-cycle representation of x . Select $t \in S_n$ such that $t^{-1}xt = x^{-1}$. Then

$$[(x_1 \cdots x_{2d})t]^{-1}x[(x_1 \cdots x_{2d})t] = t^{-1}(x_{2d} \cdots x_1)x(x_1 \cdots x_{2d})t = t^{-1}xt = x^{-1};$$

and clearly precisely one of the two elements t and

$$(x_1 \cdots x_{2d})t = (x_1x_2)(x_1x_3) \cdots (x_1x_{2d})t$$

is in A_n . But that makes x ambivalent in A_n . Thus for x to be nonambivalent, each I_i must be odd.

Now assume x is nonambivalent but Restriction 2 is not satisfied. Let $(x_1x_2 \cdots x_w)$ and $(y_1y_2 \cdots y_w)$ be two different cycles of the same order in x 's representation. From the above, w is odd. Selecting any $t \in S_n$ such that $t^{-1}xt = x^{-1}$, then

$$[(x_1y_1) \cdots (x_wy_w)t]^{-1}x[(x_1y_1) \cdots (x_wy_w)t] = t^{-1}xt = x^{-1}.$$

Again, precisely one of t and $(x_1y_1) \cdots (x_wy_w)t$ is in A_n , making x ambivalent in A_n . Thus Restriction 2 must be satisfied whenever x is nonambivalent.

(b) Finally, let $x = (x_{1,1}x_{1,2} \cdots x_{1,I_1})(x_{2,1} \cdots x_{2,I_2}) \cdots (x_{m,1} \cdots x_{m,I_m})$ be any element satisfying the first two restrictions. Then

$$x^{-1} = (x_{m,I_m}x_{m,I_m-1} \cdots x_{m,1}) \cdots (x_{1,I_1} \cdots x_{1,1}).$$

Since for any cycle y and any element $t \in S_n$, $t^{-1}yt$ is a cycle of the same length as y [2, p. 76] and since

$$t^{-1}xt = t^{-1}(x_{1,1} \cdots x_{1,I_1})t \cdot t^{-1}(x_{2,1} \cdots x_{2,I_2})t \cdots t^{-1}(x_{m,1} \cdots x_{m,I_m})t,$$

Restriction 2 necessitates that any conjugating element t sending $x \rightarrow x^{-1}$ must take each of the m cycles to its inverse. By Restriction 1, each of the elements of S_{I_i} conjugating $(x_{i,1} \cdots x_{i,I_i})$ to its inverse has a disjoint-cycle representation consisting of precisely $(I_i - 1)/2$ transpositions. There are I_i such conjugating elements in S_{I_i} . This yields $\prod_{i=1}^m I_i$ conjugating elements in S_n taking $x \rightarrow x^{-1}$, each having a disjoint-cycle representation with $\sum_{i=1}^m \frac{1}{2}(I_i - 1) = \frac{1}{2}(n - m)$ transpositions. These $\prod_{i=1}^m I_i$ elements exhaust all such conjugating elements since $o(S_n)/\prod_{i=1}^m I_i$ is indeed the order of x 's conjugate class in S_n . Thus either $\frac{1}{2}(n - m)$ is even and all $\prod_{i=1}^m I_i$ conjugating elements taking $x \rightarrow x^{-1}$ are in A_n , or $\frac{1}{2}(n - m)$ is odd and none of the conjugating elements is in A_n . In the first case x is ambivalent, in the second x is nonambivalent.

THEOREM 2. *All alternating symmetric groups A_n are nonambivalent except $A_1, A_2, A_5, A_6, A_{10}, A_{14}$.*

Proof. **Case 1:** $n = 4t$ where t is an arbitrary positive integer. Here the element $(x_1 \cdots x_{4t-1})$ is contained in A_{4t} but is nonambivalent by Theorem 1, its cycle decomposition being $\{4t - 1, 1\}$ and $\frac{1}{2}(n - m)$ being $\frac{1}{2}(4t - 2) = 2t - 1$.

Case 2: $n = 4t - 1$. As in Case 1, $(x_1 \cdots x_{4t-1})$ is in A_{4t-1} and is nonambivalent, with $\frac{1}{2}(n - m) = \frac{1}{2}((4t - 1) - 1) = 2t - 1$.

Case 3: $n = 4t + 1$. The element $(x_1 x_2 x_3) (x_4 \cdots x_{4t}) \in A_{4t+1}$ is nonambivalent provided the cycle $(x_4 \cdots x_{4t})$ has length greater than 3. Thus A_{4t+1} is nonambivalent for $t > 1$.

Case 4: $n = 4t + 2$. Here $(x_1 x_2 x_3) (x_4 \cdots x_8) (x_9 \cdots x_{4t+1})$ is nonambivalent provided $4t + 1 \geq 15$.

The above four cases show that all A_n with $n \neq 1, 2, 5, 6, 10$, or 14 are nonambivalent. The ambivalence of the remaining groups follows a quick survey that no element of those groups can satisfy the restrictions of Theorem 1.

References

1. Robert Higgins and David Ballew, An equation for finite groups, this MONTHLY, 78 (1971) 274-275.
2. I. N. Herstein, Topics in Algebra, Blaisdell, New York, 1965.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

CAN $\phi(n)$ PROPERLY DIVIDE $n - 1$?

RONALD ALTER, University of Kentucky

The ϕ in the title is the Euler function $\phi(x)$ denoting the number of natural numbers $\leq x$ which are relatively prime to x . It is well known and easy to show that $\phi(n) = n - 1$ if and only if n is a prime. The question of whether or not there exists a composite integer n and an integer $k > 1$ so that the equation

$$(1) \quad k \phi(n) = n - 1$$

has a solution, was first raised by D. H. Lehmer [2]. Lehmer proved that if such an n exists then it must be odd, square free, and the product of at least seven distinct prime numbers. Bateman and Kohlbecker [1, p. 238, exercise 14] raise the question again when they state that there are no known examples for which m is not a prime and $\phi(m) \mid (m - 1)$. Marshall [4] asks for a proof that equation (1) has no solutions for any natural number k , where n is an odd square-free natural number such that no prime factor divides the Euler function of any other prime factor.

Lehmer [2] also proves the following two theorems.

THEOREM 1. *If p is a factor of n , then n contains no prime factors of the form $px + 1$.*

THEOREM 2. *If n is a solution of equation (1) for $k = 3$, then n is a product of more than 32 distinct prime factors.*

Schuh [5] claimed that if n is composite it consists of at least 11 distinct primes. He proved the following theorem.

THEOREM 3. *If $3 \mid n$ then k is of the form $3x + 1$.*

Recently, Lieuwens [3] extended Lehmer's main result by proving that n must be a product of at least 11 distinct odd primes. He also proved the following two theorems.

THEOREM 4. *If $3 \mid n$ then n is the product of more than 212 prime numbers and $n > 5.5 \times 10^{570}$.*

THEOREM 5. *If the smallest prime factor of n is ≥ 7 then n is the product of at least 13 primes.*

With current computing facilities one would think these results could be extended. However, a proof of the conjecture that equation (1) has no solutions for $k > 1$ still appears to be extremely difficult.

References

1. E. Landau, *Elementary Number Theory* (transl. by J. E. Goodman with exercises by P. T. Bateman and E. E. Kohlbecker), Chelsea, New York, 1958.
2. D. H. Lehmer, On Euler's totient function, *Bull. Amer. Math. Soc.*, 38(1932) 745-751.
3. E. Lieuwens, Do there exist composite numbers M for which $k\phi(M) = M - 1$ holds? *Nieuw. Arch. Wisk* (3), 18(1970) 165-169.
4. A. Marshall, Problem E 2237, this MONTHLY, 77(1970) 522.
5. F. Schuh, Can $n - 1$ be divided by $\phi(n)$ when n is composite? (Dutch), *Math. Zutphen B*, 13 (1944) 102-107.

CLASSROOM NOTES

EDITED BY ROBERT GILMER

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306. Notes are usually limited to three printed pages.

A DISCOVERY APPROACH TO e

J. P. TULL, The Ohio State University

While teaching the first year course in mathematics at the University of Zambia in 1970 and 1971, I came up with the following approach to exponential and logarithmic functions. It is perhaps novel in one small way. Namely, once we had defined the exponential and logarithmic functions to an arbitrary base, starting with the

usual $a^{p/q} = \sqrt[q]{a^p}$, we eventually came to the question of the derivative. As usual we found

$$(a^{x+h} - a^x)/h = a^x(a^h - 1)/h$$

and so we needed only find the derivative at 0.

We noticed, by looking at graphs, that the derivative at 0 increases as a increases, and there are large values and small values. There ought to be one particular base, call it e , for which this derivative is 1. (Needless to say, we knew very little about continuous functions at this stage.)

For this base e , since the graph of \log is the reflection in $y = x$ of the graph of \exp , then $\log' 1 = 1$. This means that

$$(*) \quad (1/\delta) \log(1 + \delta) \rightarrow 1$$

as $\delta \rightarrow 0$. We readily find that $\log' x = 1/x$ by the direct approach, using (*). But also from (*) we see that as $\delta \rightarrow 0$

$$e^{(1/\delta) \log(1 + \delta)} \rightarrow e^1 = e;$$

i.e., $(1 + \delta)^{1/\delta} \rightarrow e$ as $\delta \rightarrow 0$. Thus $e = \lim_{n \rightarrow \infty} (1 + 1/n)^n = \lim_{n \rightarrow \infty} (1 - 1/n)^{-n}$.

SIMPLE PROOFS OF TWO ESTIMATES FOR e

R. B. DARST, Colorado State University

Let $a_n = [1 + (1/n)]^{(n+1)}$, $b_n = [1 - (1/n)]^n$, and $c_n = [1 + (1/n)]^n$, $n = 1, 2, \dots$

In elementary calculus classes one frequently establishes (C): the sequence $\{c_n\}$ increases ($c_n < c_{n+1}$); sometimes one also shows (A): $\{a_n\}$ decreases, and (B): $\{b_n\}$ increases. Then $\lim a_n = \lim c_n$, and one can show that this common value is e .

We shall give simple proofs of (A), (B) and (C) based on the fact that $(1+x)^{(n+1)} > 1 + (n+1)x$ when $x > 0$ and n is a positive integer. Thus, to establish (A), notice that

$$\begin{aligned} (a_n/a_{n+1}) &= [(n+1)/n]^{(n+1)} / \{[(n+2)/(n+1)]^{(n+1)} [1 + 1/(n+1)]\} \\ &= [(n+1)^2/(n^2 + 2n)]^{(n+1)} / [1 + 1/(n+1)] \\ &= [1 + 1/(n^2 + 2n)]^{(n+1)} / [1 + 1/(n+1)] \\ &> [1 + 1/(n+1)^2]^{(n+1)} / [1 + 1/(n+1)] > 1. \end{aligned}$$

Since $b_{(n+1)} = 1/a_n$, (B) is also established. Finally,

$$\begin{aligned} (c_{n+1}/c_n) &= [(n+2)/(n+1)]^{n+1} / \{[(n+1)/n]^{n+1} [n/(n+1)]\} \\ &= [(n^2 + 2n)/(n+1)^2]^{n+1} / [1 - 1/(n+1)] \\ &= [1 - 1/(n+1)^2]^{n+1} / [1 - 1/(n+1)] \\ &= [b_{(n+1)^2}/b_{(n+1)}]^{1/(n+1)} > 1. \end{aligned}$$

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

ASSOCIATE EDITORS: JOSHUA BARLAZ, ERIC S. LANGFORD. COLLABORATING EDITORS: LEONARD CARLITZ, GULBANK D. CHAKERIAN, HASKELL COHEN, S. ASHBY FOOTE, ISRAEL N. HERSTEIN, MURRAY S. KLAMKIN, DANIEL J. KLEITMAN, ROGER C. LYNDON, MARVIN MARCUS, CHRISTOPH NEUGEBAUER, ALBERT WILANSKY, AND UNIVERSITY OF MAINE PROBLEMS GROUP: EARL M. L. BEARD, GEORGE S. CUNNINGHAM, CLAYTON W. DODGE, HOWARD W. EVES, OSKAR FEICHTINGER, WILLIAM R. GEIGER, GARY HAGGARD, PHILIP M. LOCKE, JOHN C. MAIRHUBER, CURTIS S. MORSE, GRATTAN P. MURPHY, EDWARD S. NORTHAM AND WILLIAM L. SOULE, JR.

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before May 31, 1973. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.

An asterisk () means neither the proposer nor the editors supplied a solution.*

E 2397.* *Proposed by K. A. Brons, Cherry Hill, New Jersey*

The ellipse has the property that corresponding to any point on it there exist two other points on it such that the tangent to the curve at any of the three points is parallel to the chord joining the other two. Do any other simple closed convex planar curves enjoy this property?

E 2398. *Proposed by C. W. Dodge, University of Maine at Orono*

Prove that the point of intersection of the diagonals of a parallelogram lies on the pedal circle for any vertex with respect to the triangle formed by the other three vertices.

E 2399. *Proposed by D. E. Daykin, University of Reading, England*

Let T be an affine transformation of the plane; that is, $T\mathbf{x} = A\mathbf{x} + \mathbf{b}$, where A is a 2×2 real matrix and \mathbf{b} is a fixed vector. Characterize those affine transformations which have no fixed points and for which $TL \subseteq L$ for no line L .

E 2400. *Proposed by A. W. Walker, Toronto, Canada*

If O , H , I , r , R are the circumcenter, orthocenter, incenter, inradius and circumradius of any scalene triangle T , and P (defined as a limit point if T is right-

angled) is the orthocenter of the pedal triangle of H , then line OI divides line segment PH internally as $r:R$.

E 2401.* *Proposed by V. F. Ivanoff, San Carlos, California*

The exterior angle bisectors of a convex polygon P_0 form a polygon P_1 , whose exterior angle bisectors form a polygon P_2 , and so on. Prove that P_n approaches a regular polygon as $n \rightarrow \infty$.

E 2402. *Proposed by David Newman, Hartford, Connecticut*

Let $n > 7$ be an integer and let $N = \{2, 3, 4, \dots, 2n\}$. Show that N has precisely $n + 5$ n -element subsets S with the property that if i and j are distinct elements of S , then $i + j \notin S$.

SOLUTIONS OF ELEMENTARY PROBLEMS

A Construction without Compasses

E 2337 [1972, 180]. *Proposed by A. W. Walker, Toronto, Canada*

Show how to locate eleven coplanar points on eleven straight lines, with each point on three lines and three points on each line, using (a) straightedge and compasses; (b) straightedge only.

I. *Solution by G. B. Robinson, SUNY at New Paltz.* Solution to part (b): Let $ABCD$ be any quadrilateral, and let F be an arbitrary point on side AB . Let $E = AC \cap BD$, $H = FE \cap CD$, $L = AH \cap FD$, $J = BL \cap AD$, $G = JE \cap BC$, and $K = GD \cap FC$. These are the eleven points. When we have shown that K, E and L are collinear, then the eleven lines will be $LEK, BLJ, ALH, CKF, DKG, GEJ, ADJ, BGC, FEH, AFB$, and CHD . If we project (in 3-space) $ABCD$ into a parallelogram, then the indicated points taken in pairs are symmetric with respect to point E , so K, E , and L are collinear. Observe that we have a total of 15 lines, since BED, AEC, FLD , and BKH are also lines.

Solution to part (a): Draw a small circle in the upper left corner of the paper where it will not get in the way. Then proceed as in part (b).

II. *Remarks by the proposer.* The familiar Pascal 9_3 and Desargues 10_3 configurations naturally arouse curiosity about an 11_3 , but the only references known to me appear in *Encyk. der Math. Wiss.*, Vol. 3, Part 1-1, p. 486 (and p. 490, where it is noted that all real n_3 configurations can be constructed with straightedge and compasses, and many with straightedge alone); a French translation appears in *Encyc. des Sci. Math.*, Tom. 3, Vol. 2, pp. 153, 158. These references give no construction details. See also *Monat. für Math. u. Phys.* 6 (1895), p. 255 where, in an appendix to a paper on 12_3 configurations, diagrams are given for all the 31 possible 11_3 configurations, but with no indication of the method of construction.

Also solved by Carolyn MacDonald, and by the proposer.

A Point on a Radical Axis

E 2338 [1972, 180]. *Proposed by A. W. Walker, Toronto, Canada*

Straight lines AP , BP , CP meet the side lines BC , CA , AB of triangle ABC at points D , E , F . By Euclidean construction, locate P so that it lies on the radical axis of circles ABC and DEF .

Solution by the proposer. Let X , Y , Z be the points where circle DEF meets the lines BC , CA , AB again; then it is known [1, 2] that the lines AX , BY , CZ concur at Q , and that the areal equation of the radical axis of circles ABC and DEF is

$$\frac{x}{x_1x_2} + \frac{y}{y_1y_2} + \frac{z}{z_1z_2} = 0$$

with ABC as reference triangle and (x_1, y_1, z_1) , (x_2, y_2, z_2) as areal coordinates of P , Q . Hence P lies on this radical axis if the isotomic conjugate of Q with coordinates $(1/x_2, 1/y_2, 1/z_2)$ lies on $x + y + z = 0$, the "line at infinity." Thus we have the following construction: through A , B , C draw parallel lines meeting BC , CA , AB at U , V , W , and locate the reflections X , Y , Z of U , V , W in the midpoints of BC , CA , AB respectively; then the circle XYZ meets the lines BC , CA , AB again at D , E , F and the lines AD , BE , CF concur at a point P satisfying the required condition.

An alternative construction for P is as follows. Take a point J on the circumcircle of the triangle formed by the lines through A , B , C parallel to BC , CA , AB and locate the reflections D , E , F of the points (AJ, BC) , (BJ, CA) , (CJ, AB) in the midpoints of BC , CA , AB respectively; then the lines AD , BE , CF concur at a point P with the stated property.

The proof depends on the following two results.

(1) If the inscribed conic with center K touches the sides BC , CA , AB of triangle ABC with centroid G at D , E , F respectively, and (vectorially) $GJ = -2 \cdot GK$, then the lines AD , BE , CF concur at P , the isotomic conjugate of J (and conversely).

(2) If the centers of two conics inscribed in a triangle T are isogonal conjugates in T , the six contact points of the conics with the sides of T are concyclic (and conversely).

A proof of (2) is given in [3], and (1) is an affine-projective generalization of the special case [4] for which K , J , P are the incenter, Nagel point, and Gergonne point of triangle ABC .

Applying (1) to the above lines AX , BY , CZ , then if the isotomic conjugate of their meet Q is at infinity, so is the center of the inscribed conic XYZ (a parabola). But the isogonal transform of the "line at infinity" is the circle ABC , so it follows from (2) that the center K of the inscribed conic DEF lies on circle ABC , and therefore by another application of (1) the corresponding point J lies on the image of this circle under the homothety $(G, -2)$, justifying the second construction.

References

1. A. Cayley, *Messenger of Math.*, (2), 23 (1894) 24.
2. Tôda Ono, *Tôhoku Math. J.*, (1), 10 (1916) 31, 199.
3. R. Deaux, *Mathesis*, 63 (1954) 218.
4. R. A. Johnson, *Modern Geometry* (1929) 184, 225.

Also partially solved by Lew Kowarski. One incorrect solution was received.

A Point not on a Radical Axis

E 2339 [1972, 180]. *Proposed by A. W. Walker, Toronto, Canada*

Points D, E, F are the feet of the perpendiculars to the sides of triangle ABC from a point P ($\neq A, B$ or C) in the plane of the triangle. Prove that P cannot lie on the radical axis of circles ABC and DEF . (Cf. Problem E2338.)

Solution by the proposer. Let $a, b, c; A, B, C; R$, be the side lengths, angles, and circumradius of triangle ABC . The areal equation (with ABC as reference triangle) of the radical axis of circles ABC, DEF is [1]

$$\sum bcx_0(cy_0 + bz_0 \cos A)(bz_0 + cy_0 \cos A)x = 0,$$

where (x_0, y_0, z_0) are areal coordinates of P . If P lies on this line, its isogonal conjugate Q with areal coordinates $(a^2/x_0, b^2/y_0, c^2/z_0)$ lies on the conic with areal equation

$$\sum a(bz + cy \cos A)(cy + bz \cos A) = 0$$

which has the simpler alternative form

$$4R^2(x + y + z)^2 - (a^2yz + b^2zx + c^2xy) = 0.$$

Using a known expression [2, 3] for the distance OQ from the circumcenter O to the point $Q(x, y, z)$, this becomes

$$4R^2 - (R^2 - OQ^2) = 0, \quad OQ^2 = -3R^2,$$

so that the conic locus of Q is an imaginary circle.

References

1. John Casey, *Treatise on Analytic Geometry*, ed. 2, (1893), p. 136.
2. *Jour. de Math. Élém.*, (3), 2, (1888), p. 102.
3. *Mathesis*, 63, (1954), 120.

Doubly Stochastic Matrices

E 2340 [1972, 180]. *Proposed by Franz Hering, University of Washington*

A square matrix is *doubly-stochastic* if its entries are non-negative and if every row sum and every column sum is one. Show that every doubly-stochastic matrix (other than the one with all entries equal) contains a 2×2 submatrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

such that either $\min(a, d) > \max(b, c)$ or $\max(a, d) < \min(b, c)$.

Solution by S. S. Mitra, Wilkes College. Let a_{st} be a maximal entry of the doubly stochastic $n \times n$ matrix. Since the matrix does not have all entries equal, $a_{st} > 1/n$. Let x be the minimum of all entries that appear in either the s th row or t th column. Assume that $x = a_{sm}$ (we can give a similar argument in case $x = a_{jt}$). Since $a_{st} > 1/n$ it follows that $a_{sm} < 1/n$; that is, $a_{st} > a_{sm}$. This observation together with the fact that the sum of the entries in both the t th and m th columns is one, allows us to conclude that for some i , $a_{it} < a_{im}$. We have $a_{st} \geq a_{im} > a_{it} \geq a_{sm}$. The submatrix consisting of the above four elements has the desired property.

Also solved by David Grinstein, Joel Levy, O. P. Lossers (Netherlands), Duston Stafford, The 3-S Group of New York, R. J. Weber, and the proposer.

A Not-So-Easy Urn Problem

E 2341 [1972, 181]. *Proposed by Harry Lass, Jet Propulsion Laboratory, California Institute of Technology*

Given n urns numbered $1, 2, \dots, n$ and k objects. Suppose that each of the objects is placed at random in one of the urns. For $r = 1, 2, \dots, n$ let E_r be the event that the number of objects in the first r urns does not exceed r . Find the probability of the joint occurrence of E_1, E_2, \dots, E_n (Cf. E2252 [1971, 797].)

I. *Solution by D. M. Bloom, Brooklyn College.* Let $f(n, k)$ be the number of ways of putting k objects into n urns so that the events E_1, \dots, E_n occur jointly. Then for $n \geq 1$,

$$(*) \quad f(n, k) = \sum_{i=0}^k \binom{k}{i} f(n-1, i),$$

this equation holding for $0 \leq k \leq n$. Suppose there are i objects in the first $n-1$ urns. There are $\binom{k}{i}$ ways to choose these i objects, and given this there are $f(n-1, i)$ ways for them to be distributed in the first $n-1$ urns so that E_1, E_2, \dots, E_{n-1} occur

jointly. The other $k-i$ objects go into the n th urn and the event E_n then occurs automatically. Using (*) and a straightforward induction on n , one can then establish that

$$f(n, k) = (n+1)^{k-1}(n+1-k)$$

if $0 \leq k \leq n$ while $f(n, k) = 0$ if $k > n$. The desired probability is then $n^{-k}f(n, k)$.

II. *Solution by P. J. Burke, Bell Telephone Laboratories.* If $k > n$, the probability is obviously zero, so that we can assume that $k \leq n$. The procedure described is equivalent to the one in which the k objects are placed in $n+1$ urns, under the condition that the first urn is left empty. Using this equivalent procedure with $n+1$ urns, let S_r be the event that the number of objects in the first r urns is less than r ; the required probability is equal to that of the joint event $S_1 \cap S_2 \cap \dots \cap S_{n+1}$ given S_1 . But the probability of S_1 is $n^k/(n+1)^k$ and the probability of $S_1 \cap S_2 \cap \dots \cap S_{n+1}$ is $(n+1-k)/(n+1)$ by Theorem 1, p. 10 of L. Takács, *Combinatorial Methods in the Theory of Stochastic Processes*. Hence the required probability is $(n+1-k)(n+1)^{k-1}/n^k$.

Also solved by R. J. Weber, and the proposer.

Editor's comment. If k is taken as a fixed percentage of n , say $k = \alpha n$, then the probability approaches $(1-\alpha)e^\alpha$ as $n \rightarrow \infty$.

Mahler's Second Congruence, Resurrected

E 2342 [1972, 181]. *Proposed (independently) by Joe Buhler, Reed College, and by M. B. Nathanson, University of Rochester*

If k and n are positive integers, what is the highest power of 2 that divides $k^n - 1$? In particular, for a fixed k , find all values of n for which $k^n \equiv 1 \pmod{2^n}$.

Solution by Wells Johnson, Bowdoin College. Let $f(k, n)$ be the highest power of 2 that divides $k^n - 1$ and let $g(k, n)$ be the highest power of 2 that divides $k^n + 1$. If k is even, then obviously $f(k, n) = 0$ so we assume that k is odd in what follows. Write $k-1 = 2^r u$ and $k+1 = 2^s v$, where $r, s \geq 1$ and u and v are odd. Then $k^n = (1+2^r u)^n = 1 + 2^r n u + 2^{2r} w$ for some integer w (possibly 0 if $n=1$). If n is odd, it follows that $f(k, n) = r$, and hence $k^n \equiv 1 \pmod{2^n}$ for all odd n which do not exceed r . A similar argument shows that $g(k, n) = s$ when n is odd.

To attack the case of even n , we note first that since $k^{2^n} - 1 = (k^n - 1)(k^n + 1)$, we have the general formula $f(k, 2n) = f(k, n) + g(k, n)$. Also, if $f(k, n) \geq 2$, then $k^n \equiv 1 \pmod{4}$ so that $g(k, n) = 1$ and hence in this case $f(k, 2n) = f(k, n) + 1$. Suppose now that $n = 2^t m$ where m is odd and $t \geq 1$. Then $f(k, 2m) = f(k, m)$

$+ g(k, m) = r + s \geq 2$ so that by an easy induction it follows that $f(k, n) = f(k, 2^t m) = f(k, 2^{t-1} m) + 1 = r + s + t - 1$. This solves the first part of the problem.

Let k be a fixed odd integer. Then every odd n which does not exceed r is a solution of the congruence $k^n \equiv 1 \pmod{2^n}$; if n is even, then n solves the congruence if and only if $r + s + t - 1 \geq n$, or equivalently if and only if $n - t \leq r + s - 1$. It is easy to see that given k (and thus r and s) there are only finitely many choices for t and m and thus for n . For example, if $k = 2^{12} - 1 = 4095$, we see that $r = 1$, and $s = 12$ and thus $4095^n \equiv 1 \pmod{2^n}$ only for $n = 1, 2, 4, 6, 8, 10, 12$ and 16 . On the other hand, if $k = 4097$ so that $r = 12$ and $s = 1$, then the congruence $4097^n \equiv 1 \pmod{2^n}$ has the solutions $1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12$, and 16 .

Also solved by Joe Albree, The Bennett College Team, M. T. Bird, D. M. Bloom, M. S. Demos, O. H. Fraser, John Goth, M. G. Greening (Australia), Lew Kowarski, J. R. Kuttler, O. P. Lossers (Netherlands), Carolyn MacDonald, Grattan Murphy, Michael Shimshoni (Israel), Edith V. Sloan, D. P. Sumner, S. J. Tillman, R. J. Weber, J. R. Weiss, Brian Wesselink, Charles Wexler, and the proposers.

Editor's comment. Nathanson comments that this problem is a special case of a theorem proved by him in *An exponential congruence of Mahler*, this MONTHLY, 79 (1972) 55–57. At the time he submitted the paper, he had forgotten about his previous problem, which has, all told, led a very hard life.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers—The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before May 31, 1973. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed, stamped postcards.

An asterisk () means neither the proposer nor the editors supplied a solution.*

5895. *Proposed by Frank Bernhart, Kansas State University*

For $n \geq 1$ distinguish n points in the interior of a plane circle C and $n + 2$ points on the circumference. The $2n + 2$ points are to be connected in pairs by $2n + 1$ noncrossing arcs within C so that (a) each point on the circumference is the endpoint of one arc, (b) each interior point is the endpoint of three arcs, and (c) each pair of endpoints on the circumference is connected by a path. In graph theory terms, a cubic tree is inscribed in a circle. Case $n = 1$ is illustrated by three radii. If the points on the circumference are labeled in cyclic order x_1, x_2, \dots, x_{n+2} , put m_i = the number of interior points on the unique path between x_i and x_{i+1} ($x_{n+3} = x_1$).

1. Find a recursive definition for the set of possible sequences

$$M = (m_1, m_2, \dots, m_{n+2}).$$

2. Start with $n = 1$ and increase the tree by successive steps each consisting of

randomly selecting a point x_i , moving it inside C , and joining it to the circumference by two new arcs. Show that for a fixed integer $k \geq 1$, its fraction of occurrences in the sequences obtained by such constructing is asymptotic to $2^{k-1}/3^k$ as $n \rightarrow \infty$.

3. Find a nonrecursive test for determining if a sequence M is a possible sequence.

5896. *Proposed by A. W. Schurle, University of North Carolina at Charlotte*

It is easy to show that the metric space (X, d) is complete if it is uniformly locally compact, i.e., if there is a positive ε such that $\{y: d(x, y) \leq \varepsilon\}$ is compact for all x . Is the converse true for the real line, i.e., is every complete metric that yields the usual topology on the line uniformly locally compact?

5897*. *Proposed by I. J. Good, Virginia Polytechnic Institute and State University*

Prove that if z is real or complex and is not zero, then

$$\begin{aligned} e^{1/z} &= 1 + \frac{1}{z-1} + \frac{1}{1} + \frac{1}{1} + \frac{1}{3z-1} + \frac{1}{1} + \frac{1}{1} + \frac{1}{5z-1} + \cdots \\ &= [1, (2n+1)z-1, 1]_{n=0}^{\infty} \text{ (in a standard self-explanatory notation)} \\ &= [1, (6n+1)z-\frac{1}{2}, (24n+12)z, (6n+5)z-\frac{1}{2}]_{n=0}^{\infty}. \end{aligned}$$

5898*. *Proposed by Sylvester Reese, Baruch College, New York City*

Is the set of zeros of all entire functions with rational coefficients (for their Maclaurin series) the field of complex numbers?

5899. *Proposed by Joel Spencer, University of California, Los Angeles*

Professor Södre is, once again, unprepared for his Epsilon-delta topology class. He has prepared the first half of his lecture in which he proves a certain n propositions P_1, \dots, P_n equivalent. He had planned the most efficient proof, by showing $P_1 \Rightarrow P_2 \Rightarrow \dots \Rightarrow P_n \Rightarrow P_1$ (The theorems $P_i \Rightarrow P_j$ take an equal amount of time to prove.) Then he notices he may essentially double the length of his proof (from n to $2n-2$) by showing $P_1 \Leftrightarrow P_2 \Leftrightarrow \dots \Leftrightarrow P_n$. This method of proof is irredundant, that is, if any implication is deleted we may not deduce that P_1, \dots, P_n are equivalent. Prove that this is the longest (in terms of number of implications) irredundant method of proof.

SOLUTIONS OF ADVANCED PROBLEMS

The Distribution of Lebesgue-measurable Sets

5821 [1971, 1027]. *Proposed by Eric Langford, University of Maine*

Let I denote the unit interval $[0, 1]$. (a) Suppose that E is a Lebesgue-measurable subset of I such that $0 < m(E) < 1$. Show that

$$(i) \inf_J \frac{m(E \cap J)}{m(J)} = 0, \quad (ii) \sup_J \frac{m(E \cap J)}{m(J)} = 1,$$

where the supremum and infimum are taken over the class of all nontrivial, proper subintervals J of I .

(b) Does there exist a set E such that for every J , $0 < m(E \cap J) < m(J)$? I.e., does there exist a set E which meets every nontrivial interval in a set of positive measure, and whose complement $I \setminus E$ does likewise?

Solution by J. W. Grossman, Massachusetts Institute of Technology.

(a) (ii). Suppose not. Then there is $\delta > 0$ such that $m(E \cap J) < (1 - \delta)m(J)$ for all J . Let $0 < \varepsilon < \delta(1 - \delta)^{-1}m(E)$ and choose an open set U containing E such that $m(U) < m(E) + \varepsilon$ (by definition of measure). Write U as the disjoint union of intervals J_i . We then have

$$m(E) + \varepsilon > m(U) = \sum m(J_i) > \frac{1}{1 - \delta} \sum m(E \cap J_i) = \frac{1}{1 - \delta} m(E)$$

or $\varepsilon > \delta(1 - \delta)^{-1}m(E)$, contrary to the choice of ε .

(i) follows by applying (ii) to the complement of E .

(b) Yes. Fix $0 < \alpha < 1$. By a generalized Cantor set on an interval J we mean the set which remains after removing at the i th stage the middle "third" of length $\alpha m(J)3^{-i}$ from each interval remaining at that stage. The measure of a generalized Cantor set is $(1 - \alpha)m(J)$, and its complement is dense in J . (See Royden, *Real Analysis*, p. 63.) Now write $I = (C_0 \cup C_1 \cup C_2 \cup \dots) \cup L$, where C_0 is the Cantor set; for $n \geq 1$, C_n is the (countable) union of generalized Cantor sets, one on each interval of $I - (C_0 \cup C_1 \cup \dots \cup C_{n-1})$; and L is what's left.

Note that for $n \geq 1$ we have $m(C_n) = \alpha^{n-1}(1 - \alpha)$. Let $E = L \cup C_1 \cup C_3 \cup C_5 \cup \dots$, so that $I - E = C_0 \cup C_2 \cup C_4 \cup \dots$. Let an interval J be given. Then some interval J' of $I - (C_0 \cup C_1 \cup C_2 \cup \dots \cup C_{2N})$ is contained in J for some large N . But then $m(J \cap E) \geq (1 - \alpha)m(J') > 0$, while $m(J \cap (I - E)) \geq \frac{1}{3}\alpha(1 - \alpha)m(J') > 0$, which implies $m(J \cap E) < m(J)$, as desired.

Also solved by Michel Bousquet, Max Broberg (Sweden), R. A. Christiansen, C. V. Coffman, R. O. Davies (England), Henry Fast, Neal Felsing, G. J. Foschini, Barbara A. Keller, Douglas Lind, S. S. Mitra, Rollin Sandberg, R. M. Warten, A. Wilansky, and the proposer.

Editorial Note. Several solvers note that (a) is an immediate consequence of the Lebesgue density theorem while (b) appears as an exercise in several texts, e.g., Natanson, *Theory of Functions of a Real Variable*, p. 88. Wilansky offers other references for part (b): A. Settari (Math. Rev. 1968. #5892), C. V. Coffman (this MONTHLY 1965, p. 941). Coffman cites a theorem of J. J. Schäffer which depends on a category argument: Let \mathbf{B} be the class of Borel sets in $[0, 1]$, and let $([0, 1], \mathbf{B}, \nu)$ be a finite measure space such that $\nu(J) > 0$ for every interval $J \subset [0, 1]$. Then there exists a Borel set D such that $\nu(J \cap D) > 0$ and $\nu(J \setminus D) > 0$ for every interval $J \subset [0, 1]$.

Number of Intersections of a Secant with its Curve

5822 [1971, 1027]. *Proposed by Joseph Malkevitch, York College, New York City*

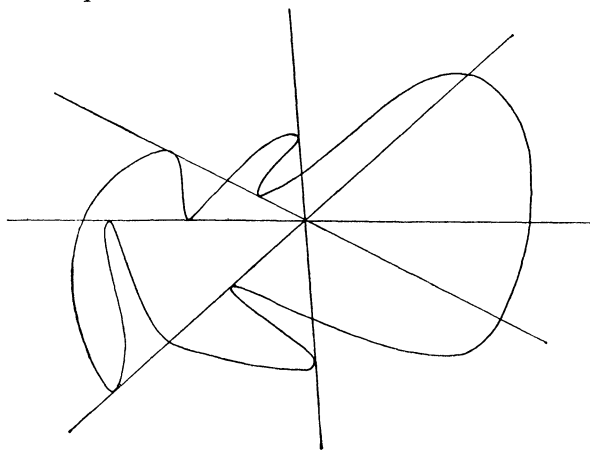
Does there exist a planar simple closed curve K such that every line through every point in the interior of K meets the boundary of K in precisely $2r$ points (r an integer ≥ 2)?

Solution by H. Guggenheimer, Polytechnic Institute of Brooklyn. The following theorem implies that the answer is in the negative.

THEOREM. *A Jordan curve of finite order (i.e., there is a finite maximal number of intersections of a line and the curve) has a secant that meets the curve in exactly two points.*

Proof. Choose any interior point P and a point X on the curve at maximal distance from P . The curve is contained in the closed disk of radius PX . Choose a secant of the circle parallel to the tangent at X and at distance ε from that tangent. The curve then intersects the circular segment bounded by the secant and the circular arc containing X in a finite number of continua. For any continuum not containing X the minimal distance from the tangent is positive. Hence, if we take ε smaller than the minimum of a finite number of positive quantities we arrive at a secant parallel to the tangent that intersects only the continuum containing X .

Since any ray issued from P intersects the curve in a finite number of points, the curve must intersect PX transversally and is starshaped for P in a neighborhood of X . X cannot be the interior point of a straight segment on the curve. A repetition of the previous argument also shows that the radius vector from P can have only a finite number of relative maxima in a neighborhood of X . Hence, the curve is locally convex near X and some secant (in fact, all secants close enough to the tangent) intersects the curve only in two points.



REMARKS. (1) The negative result also follows from Otto Haupt and Hermann Künneth, *Geometrische Ordnungen*, Springer 1967, 2. Satz, Sec. 1.5. That theorem is of extraordinary generality. The theorem proved can be obtained without effort not only for straight lines but for general order characteristics from Behauptung II, 1. Satz, same reference.

(2) The adjoining figure gives a curve that admits one point such that all lines through that point intersect the curve in 4 points. Query: Does there exist a curve with two such points?

(3) Some introductions to the geometry of plane curves are: C. Juel, *Einleitung in die Theorie der ebenen Elementarkurven dritter and vierter Ordnung*, Kgl. Danske Vidensk. Selsk. Skr., (7) 11, (1914) 113–167.

L. Locher-Ernst, *Einführung in die freie Geometrie ebener Kurven*, Birkhäuser, Basel 1952.

A. Marchaud, *Sur les continus d'ordre borné*, Acta Math., 55 (1930) 67–115.

Also solved by L. E. Mattics, and by Max Onoberg.

Domains with Linearly Ordered Primary Ideals

5823 [1971, 1028]. *Proposed by J. T. Arnold, Virginia Polytechnic Institute, and J. W. Brewer, University of Kansas*

Let D be an integral domain with identity. Show that if the primary ideals of D are linearly ordered under \subseteq and if D satisfies the ascending chain condition on prime ideals, then D is a valuation ring.

Solution by Søren Jøndrup, University of Copenhagen, Denmark. Proof is to be by induction on n , the number of prime ideals in D . We shall use two lemmas:

LEMMA 1. *If the commutative ring R has only one prime ideal, then 0 is a primary ideal.*

We observe that the set of zero divisors is a union of prime ideals and that the set of nilpotent elements is an intersection of prime ideals. Thus, with only one prime ideal, every zero divisor is nilpotent.

LEMMA 2. *If a commutative ring R has linearly ordered primary ideals, then for every prime ideal P in R , the ring R_P has the same property.*

$I \rightarrow I_P$ is a one-to-one order preserving correspondence between the primary ideals of R contained in P and the primary ideals of R_P .

Let $n = 1$. Let P be the only prime ideal of D . Let I be a proper ideal of R . If $I \subseteq P$, then R/I has just one prime ideal, P/I . By Lemma 1, I is primary. If $I \not\subseteq P$, then R/I has no prime ideals, so every element of R/I is nilpotent and again I is

primary. Since the primary ideals are linearly ordered, for any $x, y \in D$ we have $xD \subseteq yD$ or $yD \subseteq xD$.

Now let D be an integral domain with exactly n prime ideals and with linearly ordered primary ideals. Let the prime ideals of D be P_1, \dots, P_n and let $P_j \subseteq P_{j+1}$ for all j . If I is a proper ideal not contained in P_{n-1} , then I is primary by Lemma 1. Since the primary ideals in D are linearly ordered, $I \supseteq P_{n-1}$. So, for any two elements x, y not both in P_{n-1} , we have $xD \subseteq yD$ or $yD \subseteq xD$. Therefore we have to prove that for $x, y \in P_{n-1}$ it is true that $xD \subseteq yD$ or $yD \subseteq xD$. If we localize at P_{n-1} , then by the induction hypothesis and Lemma 2,

$$xD_{P_{n-1}} \subseteq yD_{P_{n-1}} \text{ or } yD_{P_{n-1}} \subseteq xD_{P_{n-1}}.$$

Let us assume that the first inclusion holds, thus we can find $r, t \in D$, $t \notin P_{n-1}$ such that $tx = ry$. We know that $rD \subseteq tD$ or $tD \subseteq rD$, and the result is proved.

Also solved by John Coolidge and by the proposer.

A Gamma Function Limit

5824 [1971, 1028]. *Proposed by N. F. Neuts, Purdue University*

Show that for every finite complex number u ,

$$\lim_{n \rightarrow +\infty} \exp(-u\gamma\sqrt{n}) \cdot \Gamma^n\left(1 - \frac{u}{\sqrt{n}}\right) = \exp\left(\frac{\pi^2 u^2}{12}\right),$$

where $\Gamma(\cdot)$ is the gamma function and γ is Euler's constant.

Solution by R. G. Buschman, University of Wyoming. For $n > |u|^2$ let $z = -u/\sqrt{n}$ in the formula

$$\log \Gamma(1+z) + \gamma z = \sum_{m=2}^{\infty} (-1)^m \zeta(m) z^m / m,$$

[*Higher Transcendental Functions*, A. Erdélyi, et al, formula 1.17 (2)]. If we multiply each side of the equation by n and take exponentials we have

$$\exp(-\gamma u\sqrt{n}) \Gamma^n(1 - u/\sqrt{n}) = \exp(u^2 \zeta(2) + f(n)).$$

Since $f(n) \rightarrow 0$ as $n \rightarrow \infty$ and $\zeta(2) = \pi^2/6$, the result follows.

Also solved by J. A. Boa, D. Borwein, Robert Breusch, Paul Bugl, M. L. Glasser, A. A. Jagers (Netherlands), Václav Konečný, O. P. Lossers (Netherlands), M. H. Moore, C. C. Rousseau, O. G. Ruehr, David Shelupsky, P. H. Young, and the proposer.

Shelupsky states the limit in the following form: If $f(z_0) = 0$, $f(z) \in C^3$ in a neighborhood of z_0 , then

$$u\sqrt{n}f'(z_0) + nf(z_0 - u/\sqrt{n}) \rightarrow u^2 f''(z_0)/2.$$

Roots of a Minimal Polynomial

5825 [1971, 1028]. *Proposed by Erwin Just, Bronx Community College*

Assume that α and β are real numbers, $\beta \neq 0$, such that $\alpha + \beta i$ is a zero of $f(x)$, a cubic polynomial with rational coefficients. If $g(x)$ is the minimal polynomial (with rational coefficients) of βi , can any of the zeros of $g(x)$ be real?

Solution by Irving Gerst, State University of New York at Stony Brook. The zeros of $f(x)$ are $\gamma_1 = \alpha + \beta i$, $\gamma_2 = \alpha - \beta i$ and a real zero γ_3 . Then

$$h(x) = \prod_{\substack{i,j=1 \\ i \neq j}}^3 [x - (\gamma_i - \gamma_j)/2]$$

has βi as a zero, has no real zeros, and, by the symmetric function theorem, has rational coefficients. Since $g(x)$ is a divisor of $h(x)$, no zero of $g(x)$ can be real.

Also solved by Robert Breusch, R. O. Davies (England), G. J. Foschini, Anne Grams & Tom Parker, P. R. Hafner (New Zealand), A. A. Jagers (Netherlands), L. E. Mattics, P. L. Montgomery, P. J. Owens (England), Nicholas Passell, Stephen Pierce, Wesley Tom, Elizabeth Yip, and the proposer.

Montgomery offers the polynomial $x^4 - 2$ to show that some restriction on the degree of $f(x)$ is necessary.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN

with the assistance of the mathematics departments of St. Olaf and Carleton Colleges

COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER, CARLETON COLLEGE

Printed materials for reviews should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057. Films and correspondence relating to films should be sent to Seymour Schuster, Carleton College, Northfield MN 55057.

All unsigned material is written by the editors. A boldface capital C in the margin indicates that a review is based in part on classroom use. Professors willing to write such a review should inform the editor in order to avoid duplication.

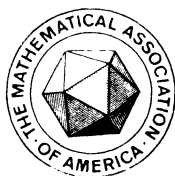
Probability and Mathematical Statistics: An Introduction. By Eugene Lukacs. Academic Press, New York, 1972. x+242 pp. \$8.50. (Telegraphic Review, April 1972.)
Introductory Statistics and Probability: A Basis for Decision Making. By David W. Blakeslee, William G. Chinn. Houghton Mifflin, Boston, Massachusetts, 1971. ix + 356 pp. \$7.95. (Telegraphic Review, April 1972.)

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF
THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 80



NUMBER 3

CODEN: AMMYAE

CONTENTS

Hilbert's Tenth Problem is Unsolvable	MARTIN DAVIS	233
History of the Riemann Mapping Theorem	J. L. WALSH	270
The First U. S. A. Mathematical Olympiad	S. GREITZER	276
Correction to "What is a Reciprocity Law?"	B. F. WYMAN	281

MATHEMATICAL NOTES

On Elementary Proofs of Peano's Existence Theorems	JOHANN WALTER	282
A Remark Concerning Absolutely Continuous Functions.	F. S. VAN VLECK	286
On Non-Associative Algebras Derived from Graphs	W. E. JENNER	288
A Finite Difference Proof that $E = mc^2$	DONALD GREENSPAN	289

RESEARCH PROBLEMS

Reachability Problems in Vector Addition Systems	B. O. NASH	292
When do All k -Sequences Modulo m have Period One?	E. A. PARBERRY AND NANCY GRAUDONS	295

CLASSROOM NOTES

On Injective Modules	AZMI HANNA	297
The Hamel Dimension of any Infinite Dimensional Separable Banach Space is c	H. ELTON LACEY	298
A Note on Conformality	R. K. WILLIAMS	299
A Wronskian Condition Related to Ordinary Differential Equations	L. C. EGGAN AND A. J. INSEL	300

MATHEMATICAL EDUCATION

Teaching Applicable Mathematics	E. A. BENDER	302
-------------------------------------------	--------------	-----

(Continued on inside cover)

MARCH

1973

Individualized Instruction in Large Enrollment Mathematics Courses .	BERT WAITS	307
Female Mathematicians, Where are You?	VIOLET H. LARNEY	310
CUPM Report to the Board of Governors, August 1972		313
ELEMENTARY PROBLEMS AND SOLUTIONS		315
ADVANCED PROBLEMS AND SOLUTIONS		324
REVIEWS		330
NEWS AND NOTICES		346
MATHEMATICAL ASSOCIATION OF AMERICA		347
Committee on Educational Media.		347
Proceedings of the 1971 Summer Conference held at the University of Missouri, Rolla .		347
Calendars of Future Meetings		348

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss.

Backlog: Main Articles 12 months, Math. Notes 13 months, Research Problems 7 months, Classroom Notes 10 months, Math. Education 10 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to ALEX ROSENBERG, Department of Mathematics, Cornell University, Ithaca, N.Y. 14850; NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D. C. 20036.

HARLEY FLANDERS, *Editor*
ALEX ROSENBERG, *Editor-Elect*

ASSOCIATE EDITORS

JOSHUA BARLAZ	J. G. HARVEY	SEYMOUR SCHUSTER
E. R. BERLEKAMP	ERIC S. LANGFORD	J. ARTHUR SEEBACH, Jr.
JANE W. DI PAOLA	P. D. LAX	E. P. STARKE
ROBERT GILMER	ARTHUR MATTUCK	LYNN A. STEEN
RICHARD GUY	M. W. POWNALL	JAMES WENDEL
RAOUL HAILPERN	GIAN-CARLO ROTA	

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are \$12.50. For nonmembers the subscription price is \$18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices.

Copyright © The Mathematical Association of America (Incorporated), 1973

PRINTED IN THE UNITED STATES OF AMERICA

HILBERT'S TENTH PROBLEM IS UNSOLVABLE

MARTIN DAVIS, Courant Institute of Mathematical Science

When a long outstanding problem is finally solved, every mathematician would like to share in the pleasure of discovery by following for himself what has been done. But too often he is stymied by the abstruseness of so much of contemporary mathematics. The recent negative solution to Hilbert's tenth problem given by Matiyasevič (cf. [23], [24]) is a happy counterexample. In this article, a complete account of this solution is given; the only knowledge a reader needs to follow the argument is a little number theory: specifically basic information about divisibility of positive integers and linear congruences. (The material in Chapter 1 and the first three sections of Chapter 2 of [25] more than suffices.)

Hilbert's tenth problem is to give a computing algorithm which will tell of a given polynomial Diophantine equation with integer coefficients whether or not it has a solution in integers. Matiyasevič proved that *there is no such algorithm*.

Hilbert's tenth problem is the tenth in the famous list which Hilbert gave in his 1900 address before the International Congress of Mathematicians (cf. [18]). The way in which the problem has been resolved is very much in the spirit of Hilbert's address in which he spoke of the conviction among mathematicians "that every definite mathematical problem must necessarily be susceptible of a precise settlement, either in the form of an actual answer to the question asked, or by *the proof of the impossibility of its solution ...*" (italics added). Concerning such impossibility proofs Hilbert commented:

"Sometimes it happens that we seek the solution under unsatisfied hypotheses or in an inappropriate sense and are therefore unable to reach our goal. Then the task arises of proving the impossibility of solving the problem under the given hypotheses and in the sense required. Such impossibility proofs were already given by the ancients, in showing, e.g., that the hypotenuse of an isosceles right triangle has an irrational ratio to its leg. In modern mathematics the question of the impossibility of certain solutions has played a key role, so that we have acquired the knowledge that such old and difficult problems as to prove the parallel axiom, to square the circle, or to solve equations of the fifth degree in radicals have no solution in the originally intended sense, but nevertheless have been solved in a precise and completely satisfactory way."

Martin Davis received his Princeton Ph. D. under Alonzo Church. He has held positions at Univ. of Illinois, IAS, Univ. of Calif.-Davis, Ohio State Univ., Rensselaer Poly, Yeshiva Univ. and New York Univ., and he spent a leave at Westfield College, London. He has done research in various aspects of the foundations of mathematics, and is the author of *Computability and Unsolvability* (McGraw-Hill, 1958), *The Undecidable* (editor, Raven Press, 1965), *Lectures on Modern Mathematics* (Gordon and Breach, 1967), and *First Course in Functional Analysis* (Gordon and Breach, 1967).
Editor.

Matiyasevič's negative solution of Hilbert's tenth problem is of just this character. It is not a solution in Hilbert's "originally intended sense" but rather a "precise and completely satisfactory" proof that no such solution is possible. The methods needed to make it possible to prove the non-existence of algorithms had not been developed in 1900. These methods are part of the theory of recursive (or computable) functions, developed by logicians much later ([6] is an exposition of recursive function theory). In this article no previous knowledge of recursive function theory is assumed. The little that is needed is developed in the article itself.

What will be proved in the body of this article is that no algorithm exists for testing a polynomial with integer coefficients to determine whether or not it has *positive integer* solutions (Hilbert inquired about arbitrary integer solutions). But then it will follow at once that there can be no algorithm for integer solutions either. For one could test the equation

$$P(x_1, \dots, x_n) = 0$$

for possession of positive solutions $\langle x_1, \dots, x_n \rangle$ by testing

$$P(1 + p_1^2 + q_1^2 + r_1^2 + s_1^2, \dots, 1 + p_n^2 + q_n^2 + r_n^2 + s_n^2) = 0$$

for possession of integer solutions $\langle p_1, q_1, r_1, s_1, \dots, p_n, q_n, r_n, s_n \rangle$. This is because (by a well-known theorem of Lagrange) every non-negative integer is the sum of four squares. (Just this once the stated prerequisite is exceeded! Cf. [17], p. 302.) In the body of this article, only positive integers will be dealt with—except when the contrary is explicitly stated.

When Matiyasevič announced his beautiful and ingenious solution in January 1970, it had been known for a decade that the unsolvability of Hilbert's tenth problem would follow if one could construct a Diophantine equation whose solutions were such that one of its components grew roughly exponentially with another of its components. (In §9, this is explained more precisely.) Matiyasevič showed how the Fibonacci numbers could be used to construct such an equation. In this article the historical development of the subject will not be followed; the aim has rather been to give as smooth and straightforward an account of the main results as seems currently feasible. A brief appendix gives the history.

1. Diophantine Sets. In this article the usual problem of Diophantine equations will be inverted. Instead of being given an equation and seeking its solutions, one will begin with the set of "solutions" and seek a corresponding Diophantine equation. More precisely:

DEFINITION. A set S of ordered n -tuples of positive integers is called **Diophantine** if there is a polynomial $P(x_1, \dots, x_n, y_1, \dots, y_m)$, where $m \geq 0$, with integer coefficients such that a given n -tuple $\langle x_1, \dots, x_n \rangle$ belongs to S if and only if there exist positive integers y_1, \dots, y_m for which

$$P(x_1, \dots, x_n, y_1, \dots, y_m) = 0.$$

Borrowing from logic the symbols “ \exists ” for “there exists” and “ \Leftrightarrow ” for “if and only if”, the relation between the set S and the polynomial P can be written succinctly as:

$$\langle x_1, \dots, x_n \rangle \in S \Leftrightarrow (\exists y_1, \dots, y_m) [P(x_1, \dots, x_n, y_1, \dots, y_m) = 0],$$

or equivalently:

$$S = \{ \langle x_1, \dots, x_n \rangle \mid (\exists y_1, \dots, y_m) [P(x_1, \dots, x_n, y_1, \dots, y_m) = 0] \}.$$

Note that P may (and in non-trivial cases always will) have negative coefficients. The word “**polynomial**” should always be so construed in the article except where the contrary is explicitly stated. Also all numbers in this article are positive integers unless the contrary is stated.

The main question which will be discussed (and settled) in this article is:

Which sets are Diophantine? A vague paraphrase of the eventual answer is: *any set which could possibly be Diophantine is Diophantine*. What does the phrase “which could possibly be Diophantine” mean? And how is all this related to Hilbert’s tenth problem? These quite reasonable questions will only be answered much later. In the meantime, the task will be developing techniques for showing that various sets are indeed Diophantine.

A few very simple examples:

(i) *the numbers which are not powers of 2:*

$$x \in S \Leftrightarrow (\exists y, z) [x = y(2z + 1)],$$

(ii) *the composite numbers:*

$$x \in S \Leftrightarrow (\exists y, z) [x = (y + 1)(z + 1)],$$

(iii) *the ordering relation on the positive integers; that is the sets $\{ \langle x, y \rangle \mid x < y \}$, $\{ \langle x, y \rangle \mid x \leq y \}$:*

$$x < y \Leftrightarrow (\exists z) (x + z = y),$$

$$x \leq y \Leftrightarrow (\exists z) (x + z - 1 = y),$$

(iv) *the divisibility relation; that is $\{ \langle x, y \rangle \mid x \mid y \}$:*

$$x \mid y \Leftrightarrow (\exists z) (xz = y).$$

Examples (i) and (ii) suggest, as other sets to consider, the set of powers of 2 and of primes respectively. As we shall eventually see, these sets are Diophantine; but the proof is not at all easy.

Another example:

(v) *the set W of $\langle x, y, z \rangle$ for which $x \mid y$ and $x < z$:* Here

$$x \mid y \Leftrightarrow (\exists u) (y = xu) \text{ and } x < z \Leftrightarrow (\exists v) (z = x + v).$$

Hence,

$$\langle x, y, z \rangle \in W \Leftrightarrow (\exists u, v) [(y - xu)^2 + (z - x - v)^2 = 0].$$

Note that the technique just used is perfectly general. So, in defining a Diophantine set one may use a *simultaneous system* $P_1 = 0, P_2 = 0, \dots, P_k = 0$ of polynomial equations since this system can be replaced by the equivalent single equation:

$$P_1^2 + P_2^2 + \dots + P_k^2 = 0.$$

By a “function” a positive integer valued function of one or more positive integer arguments will always be understood.

DEFINITION. A function f of n arguments is called **Diophantine** if

$$\{\langle x_1, \dots, x_n, y \rangle \mid y = f(x_1, \dots, x_n)\}$$

is a Diophantine set, (i.e., f is Diophantine if its “graph” is Diophantine).

Another question that will be answered here is: *which functions are Diophantine?*

An important Diophantine function is associated with the *triangular numbers*, that is numbers of the form:

$$T(n) = 1 + 2 + \dots + n = \frac{n(n+1)}{2}.$$

Since $T(n)$ is an increasing function, for each positive integer z , there is a unique $n \geq 0$ such that

$$T(n) < z \leq T(n+1) = T(n) + n + 1.$$

Hence each z is *uniquely representable* as:

$$z = T(n) + y; \quad y \leq n + 1,$$

or equivalently, uniquely representable as:

$$z = T(x + y - 2) + y.$$

In this case, one writes $x = L(z)$, $y = R(z)$; also one sets

$$P(x, y) = T(x + y - 2) + y - 1.$$

Note that $L(z)$, $R(z)$ and $P(x, y)$ are Diophantine functions since

$$z = P(x, y) \Leftrightarrow 2z = (x + y - 2)(x + y - 1) + 2y$$

$$x = L(z) \Leftrightarrow (\exists y) [2z = (x + y - 2)(x + y - 1) + 2y]$$

$$y = R(z) \Leftrightarrow (\exists x) [2z = (x + y - 2)(x + y - 1) + 2y].$$

The function $P(x, y)$ maps the set of ordered pairs of positive integers one-one

onto the set of positive integers. And, for each z , the ordered pair which is mapped into z by $P(x, y)$ is $(L(z), R(z))$. ("P" is for "pair", "L" for "left", and "R" for "right".) Note also that $L(z) \leq z$, $R(z) \leq z$. To summarize:

THEOREM 1.1 (Pairing Function Theorem¹). *There are Diophantine functions $P(x, y)$, $L(z)$, $R(z)$ such that*

- (1) *for all x, y , $L(P(x, y)) = x$, $R(P(x, y)) = y$, and*
- (2) *for all z , $P(L(z), R(z)) = z$, $L(z) \leq z$, $R(z) \leq z$.*

Another useful Diophantine function is related to the Chinese Remainder Theorem, stated below:

DEFINITION. The numbers m_1, \dots, m_N are called an **admissible sequence of moduli** if $i \neq j$ implies that m_i and m_j are relatively prime.

THEOREM 1.2 (Chinese Remainder Theorem). *Let a_1, \dots, a_N be any positive integers and let m_1, \dots, m_N be an admissible sequence of moduli. Then there is an x such that:*

$$\begin{aligned} x &\equiv a_1 \pmod{m_1} \\ x &\equiv a_2 \pmod{m_2} \\ &\dots \dots \dots \\ x &\equiv a_N \pmod{m_N}. \end{aligned}$$

The Chinese remainder theorem is proved for example in [25], p. 33. (That x can be assumed positive is not ordinarily stated. But since the product of the moduli added to a solution gives another solution, this is obvious.)

Now let the function $S(i, u)$ be defined as follows:

$$S(i, u) = w,$$

where w is the unique positive integer for which:

$$w \equiv L(u) \pmod{1 + iR(u)}$$

$$w \leq 1 + iR(u).$$

Here w is simply the least *positive* remainder when $L(u)$ is divided by $1 + iR(u)$.

THEOREM 1.3 (Sequence Number Theorem). *There is a Diophantine function $S(i, u)$ such that*

- (1) *$S(i, u) \leq u$, and*
- (2) *for each sequence a_1, \dots, a_N , there is a number u such that*

$$S(i, u) = a_i \text{ for } 1 \leq i \leq N.$$

Proof. The first task is to show that $S(i, u)$ as defined just above, is a Diophantine

function. The claim is that $w = S(i, u)$ if and only if the following system of equations has a solution:

$$2u = (x + y - 2)(x + y - 1) + 2y$$

$$x = w + z(1 + iy)$$

$$1 + iy = w + v - 1.$$

This is because (by the discussion leading to the Pairing Function Theorem), the first equation is equivalent to:

$$x = L(u) \text{ and } y = R(u).$$

Then (using a technique already noted) one needs only sum the squares of the three equations to see that $S(i, u)$ is Diophantine.

Now $S(i, u) \leq L(u) \leq u$. So finally, let a_1, \dots, a_N be given numbers. Choose y to be some number greater than each of a_1, \dots, a_N and divisible by each of $1, 2, \dots, N$. Then the numbers $1 + y, 1 + 2y, \dots, 1 + Ny$ are an admissible sequence of moduli. (For, if $d \mid 1 + iy$ and $d \mid 1 + jy$, $i < j$, then $d \mid [j(1 + iy) - i(1 + jy)]$, i.e., $d \mid j - i$ so that $d \leq N$; but this is impossible unless $d = 1$ because $d \mid y$.) This being the case, the Chinese Remainder Theorem can be applied to obtain a number x such that

$$x \equiv a_1 \pmod{1 + y}$$

$$x \equiv a_2 \pmod{1 + 2y}$$

$$\dots \dots \dots$$

$$x \equiv a_N \pmod{1 + Ny}.$$

Let $u = P(x, y)$, so that $x = L(u)$ and $y = R(u)$. Then, for $i = 1, 2, \dots, N$

$$a_i \equiv L(u) \pmod{1 + iR(u)}$$

and $a_i < y = R(u) < 1 + iR(u)$. But then by definition, $a_i = S(i, u)$.

A striking characterization of Diophantine sets of positive integers (cf. [26]) is given by:

THEOREM 1.4. *A set S of positive integers is Diophantine if and only if there is a polynomial P such that S is precisely the set of positive integers in the range of P .*

Proof. If S is related to $P(x_1, \dots, x_m)$ as in the theorem then

$$x \in S \Leftrightarrow (\exists x_1, \dots, x_m) [x = P(x_1, \dots, x_m)].$$

Conversely, let

$$x \in S \Leftrightarrow (\exists x_1, \dots, x_m) [Q(x, x_1, \dots, x_m) = 0].$$

Let $P(x, x_1, \dots, x_m) = x[1 - Q^2(x, x_1, \dots, x_m)]$. Then, if $x \in S$, choose x_1, \dots, x_m such

that $Q(x, x_1, \dots, x_m) = 0$. Then $P(x, x_1, \dots, x_m) = x$; so x is in the range of P . On the other hand, if $z = P(x, x_1, \dots, x_m)$, $z > 0$, then $Q(x, x_1, \dots, x_m)$ must vanish (otherwise $1 - Q^2 \leq 0$) so that $z = x$ and $x \in S$.

2. Twenty-four easy lemmas. The first major task is to prove that the exponential function $h(n, k) = n^k$ is Diophantine. This is the hardest thing we shall have to do. The proof is in §3. In this section we develop the methods we shall need, using the so-called Pell equation:

$$\text{where } \left. \begin{aligned} x^2 - dy^2 &= 1, & x, y &\geq 0, \\ d &= a^2 - 1, & a &> 1. \end{aligned} \right\} \quad (*)$$

Although this is a famous equation with a considerable literature,² a self-contained treatment is given. Note the obvious solutions to (*):

$$x = 1 \quad y = 0$$

$$x = a \quad y = 1.$$

LEMMA 2.1. *There are no integers x, y , positive, negative, or zero, which satisfy (*) for which $1 < x + y\sqrt{d} < a + \sqrt{d}$.*

Proof. Let x, y satisfy (*). Since

$$1 = (a + \sqrt{d})(a - \sqrt{d}) = (x + y\sqrt{d})(x - y\sqrt{d}),$$

the inequality implies (taking negative reciprocals) $-1 < -x + y\sqrt{d} < -a + \sqrt{d}$. Adding the inequalities: $0 < 2y\sqrt{d} < 2\sqrt{d}$, i.e., $0 < y < 1$, a contradiction.

LEMMA 2.2. *Let x, y and x', y' be integers, positive, negative, or zero which satisfy (*). Let*

$$x'' + y''\sqrt{d} = (x + y\sqrt{d})(x' + y'\sqrt{d}).$$

Then, x'', y'' satisfies ().*

Proof. Taking conjugates: $x'' - y''\sqrt{d} = (x - y\sqrt{d})(x' - y'\sqrt{d})$. Multiplying gives:

$$(x'')^2 - d(y'')^2 = (x^2 - dy^2)((x')^2 - d(y')^2) = 1.$$

DEFINITION. $x_n(a), y_n(a)$ are defined for $n \geq 0, a > 1$, by setting

$$x_n(a) + y_n(a)\sqrt{d} = (a + \sqrt{d})^n.$$

Where the context permits, the dependence on a is not explicitly shown, writing x_n, y_n .

LEMMA 2.3. x_n, y_n satisfy (*).

Proof. This follows at once by induction using Lemma 2.2.

LEMMA 2.4. *Let x, y be a non-negative solution of (*). Then for some n , $x = x_n$, $y = y_n$.*

Proof. To begin with $x + y\sqrt{d} \geq 1$. On the other hand the sequence $(a + \sqrt{d})^n$ increases to infinity. Hence for some $n \geq 0$,

$$(a + \sqrt{d})^n \leq x + y\sqrt{d} < (a + \sqrt{d})^{n+1}.$$

If there is equality, the result is proved; so suppose otherwise:

$$x_n + y_n\sqrt{d} < x + y\sqrt{d} < (x_n + y_n\sqrt{d})(a + \sqrt{d}).$$

Since $(x_n + y_n\sqrt{d})(x_n - y_n\sqrt{d}) = 1$, the number $x_n - y_n\sqrt{d}$ is positive. Hence, $1 < (x + y\sqrt{d})(x_n - y_n\sqrt{d}) < a + \sqrt{d}$. But this contradicts Lemmas 2.1 and 2.2.

The defining relation:

$$x_n + y_n\sqrt{d} = (a + \sqrt{d})^n$$

is a formal analogue of the familiar formula:

$$(\cos u) + (\sin u)\sqrt{-1} = e^{iu} = (\cos 1 + (\sin 1)\sqrt{-1})^u,$$

with x_n playing the role of \cos , y_n playing the role of \sin and d playing the role of -1 . Thus, the familiar trigonometric identities have analogues in which -1 is replaced by d at appropriate places. For example the Pell equation itself

$$x_n^2 - dy_n^2 = 1$$

is just the analogue of the Pythagorean identity. Next analogues of the familiar addition formulas are obtained.

LEMMA 2.5. $x_{m \pm n} = x_m x_n \pm dy_n y_m$ and $y_{m \pm n} = x_n y_m \pm x_m y_n$.

Proof.

$$\begin{aligned} x_{m+n} + y_{m+n}\sqrt{d} &= (a + \sqrt{d})^{m+n} \\ &= (x_m + y_m\sqrt{d})(x_n + y_n\sqrt{d}) \\ &= (x_m x_n + dy_n y_m) + (x_n y_m + x_m y_n)\sqrt{d}. \end{aligned}$$

Hence,

$$x_{m+n} = x_m x_n + dy_n y_m$$

$$y_{m+n} = x_n y_m + x_m y_n.$$

Similarly, $(x_{m-n} + y_{m-n}\sqrt{d})(x_n + y_n\sqrt{d}) = x_m + y_m\sqrt{d}$. So

$$x_{m-n} + y_{m-n}\sqrt{d} = (x_m + y_m\sqrt{d})(x_n - y_n\sqrt{d}),$$

and one proceeds as above.

LEMMA 2.6. $y_{m \pm 1} = a y_m \pm x_m$, and $x_{m \pm 1} = a x_m \pm d y_m$.

Proof. Take $n = 1$ in Lemma 2.5.

The familiar notation (x, y) is used to symbolize the g.c.d. of x and y .

LEMMA 2.7. $(x_n, y_n) = 1$.

Proof. If $d \mid x_n$ and $d \mid y_n$, then $d \mid x_n^2 - d y_n^2$, i.e., $d \mid 1$.

LEMMA 2.8. $y_n \mid y_{nk}$.

Proof. This is obvious when $k = 1$. Proceeding by induction, using the addition formula (Lemma 2.5),

$$y_{n(m+1)} = x_n y_{nm} + x_{nm} y_n.$$

By the induction hypothesis $y_n \mid y_{nm}$. Hence, $y_n \mid y_{n(m+1)}$.

LEMMA 2.9. $y_n \mid y_t$ if and only if $n \mid t$.

Proof. Lemma 2.8 gives the implication in one direction. For the converse suppose $y_n \mid y_t$ but $n \nmid t$. So one can write $t = nq + r$, $0 < r < n$. Then,

$$y_t = x_r y_{nq} + x_{nq} y_r.$$

Since (by Lemma 2.8) $y_n \mid y_{nq}$, it follows that $y_n \mid x_{nq} y_r$. But $(y_n, x_{nq}) = 1$. (If $d \mid y_n$, $d \mid x_{nq}$, then by Lemma 2.8 $d \mid y_{nq}$ which, by Lemma 2.7, implies $d = 1$.) Hence $y_n \mid y_r$. But, since $r < n$, we have $y_r < y_n$ (e.g., by Lemma 2.6). This is a contradiction.

LEMMA 2.10. $y_{nk} \equiv k x_n^{k-1} y_n \pmod{(y_n)^3}$.

Proof.

$$\begin{aligned} x_{nk} + y_{nk} \sqrt{d} &= (a + \sqrt{d})^{nk} \\ &= (x_n + y_n \sqrt{d})^k \\ &= \sum_{j=0}^k \binom{k}{j} x_n^{k-j} y_n^j d^{j/2}. \end{aligned}$$

So,

$$y_{nk} = \sum_{\substack{j=1 \\ j \text{ odd}}}^k \binom{k}{j} x_n^{k-j} y_n^j d^{(j-1)/2}.$$

But all terms of this expansion for which $j > 1$ are $\equiv 0 \pmod{(y_n)^3}$.

LEMMA 2.11. $y_n^2 \mid y_{ny_n}$.

Proof. Set $k = y_n$ in Lemma 2.10.

LEMMA 2.12. If $y_n^2 \mid y_t$, then $y_n \mid t$.

Proof. By Lemma 2.9, $n \mid t$. Set $t = nk$. Using Lemma 2.10, $y_n^2 \mid k x_n^{k-1} y_n$, i.e., $y_n \mid k x_n^{k-1}$. But by Lemma 2.7, $(y_n, x_n) = 1$. So, $y_n \mid k$ and hence $y_n \mid t$.

LEMMA 2.13. $x_{n+1} = 2ax_n - x_{n-1}$ and $y_{n+1} = 2ay_n - y_{n-1}$.

Proof. By Lemma 2.6,

$$\begin{aligned} x_{n+1} &= ax_n + dy_n, & y_{n+1} &= ay_n + x_n, \\ x_{n-1} &= ax_n - dy_n, & y_{n-1} &= ay_n - x_n. \end{aligned}$$

So, $x_{n+1} + x_{n-1} = 2ax_n$, $y_{n+1} + y_{n-1} = 2ay_n$.

These second order difference equations, together with the initial values $x_0 = 1$, $x_1 = a$, $y_0 = 0$, $y_1 = 1$, determine the values of all the x_n , y_n . Various properties of these sequences are easily established by checking them for $n = 0, 1$ and using these difference equations to show that the property for $n + 1$ can be inferred from its holding for n and $n - 1$. Some simple (but important) examples follow:

LEMMA 2.14. $y_n \equiv n \pmod{a-1}$.

Proof. For $n = 0, 1$ equality holds. Proceeding inductively, using $a \equiv 1 \pmod{a-1}$:

$$\begin{aligned} y_{n+1} &= 2ay_n - y_{n-1} \\ &\equiv 2n - (n-1) \pmod{a-1}. \end{aligned}$$

LEMMA 2.15. If $a \equiv b \pmod{c}$, then for all n ,

$$x_n(a) \equiv x_n(b), \quad y_n(a) \equiv y_n(b) \pmod{c}.$$

Proof. Again for $n = 0, 1$ the congruence is an equality. Proceeding by induction:

$$\begin{aligned} y_{n+1}(a) &= 2ay_n(a) - y_{n-1}(a) \\ &\equiv 2by_n(b) - y_{n-1}(b) \pmod{c} \\ &= y_{n+1}(b). \end{aligned}$$

LEMMA 2.16. When n is even y_n is even and when n is odd y_n is odd.

Proof. $y_{n+1} = 2ay_n - y_{n-1} \equiv y_{n-1} \pmod{2}$. So when n is even, $y_n \equiv y_0 = 0 \pmod{2}$, and when n is odd, $y_n \equiv y_1 = 1 \pmod{2}$.

LEMMA 2.17. $x_n(a) - y_n(a)(a - y) \equiv y^n \pmod{2ay - y^2 - 1}$.

Proof. $x_0 - y_0(a - y) = 1$ and $x_1 - y_1(a - y) = y$, so the result holds for $n = 0$ and 1 . Using Lemma 2.13 and proceeding by induction:

$$\begin{aligned} x_{n+1} - y_{n+1}(a - y) &= 2a[x_n - y_n(a - y)] - [x_{n-1} - y_{n-1}(a - y)] \\ &\equiv 2ay^n - y^{n-1} \end{aligned}$$

$$\begin{aligned}
 &= y^{n-1}(2ay - 1) \\
 &\equiv y^{n-1} y^2 \\
 &= y^{n+1}.
 \end{aligned}$$

LEMMA 2.18. For all n , $y_{n+1} > y_n \geq n$.

Proof. By Lemma 2.6, $y_{n+1} > y_n$. Since $y_0 = 0 \geq 0$, it follows by induction that $y_n \geq n$ for all n .

LEMMA 2.19. For all n , $x_{n+1}(a) > x_n(a) \geq a^n$; $x_n(a) \leq (2a)^n$.

Proof. By Lemmas 2.6 and 2.13 $a x_n(a) \leq x_{n+1}(a) \leq (2a)x_n(a)$. The result follows by induction.

Next some periodicity properties of the sequence x_k are obtained.

LEMMA 2.20. $x_{2n \pm j} \equiv -x_j \pmod{x_n}$.

Proof. By the addition formulas (Lemma 2.5)

$$\begin{aligned}
 x_{2n \pm j} &= x_n x_{n \pm j} + d y_n y_{n \pm j} \\
 &\equiv d y_n (y_n x_j \pm x_n y_j) \pmod{x_n} \\
 &\equiv d y_n^2 x_j \pmod{x_n} \\
 &= (x_n^2 - 1) x_j \\
 &\equiv -x_j \pmod{x_n}.
 \end{aligned}$$

LEMMA 2.21. $x_{4n \pm j} \equiv x_j \pmod{x_n}$.

Proof. By Lemma 2.20

$$x_{4n \pm j} \equiv -x_{2n \pm j} \equiv x_j \pmod{x_n}.$$

LEMMA 2.22. Let $x_i \equiv x_j \pmod{x_n}$, $i \leq j \leq 2n$, $n > 0$. Then $i = j$, unless $a = 2$, $n = 1$, $i = 0$ and $j = 2$.

Proof. First suppose x_n is odd and let $q = (x_n - 1)/2$. Then the numbers $-q, -q + 1, -q + 2, \dots, -1, 0, 1, \dots, q - 1, q$ are a complete set of mutually incongruent residues modulo x_n . Now by Lemma 2.19,

$$1 = x_0 < x_1 < \dots < x_{n-1}.$$

Using Lemma 2.6, $x_{n-1} \leq x_n/a \leq \frac{1}{2}x_n$; so $x_{n-1} \leq q$. Also by Lemma 2.20, the numbers

$$x_{n+1}, x_{n+2}, \dots, x_{2n-1}, x_{2n}$$

are congruent modulo x_n respectively to:

$$-x_{n-1}, -x_{n-2}, \dots, -x_1, -x_0 = -1.$$

Thus the numbers $x_0, x_1, x_2, \dots, x_{2n}$ are mutually incongruent modulo x_n . This gives the result.

Next suppose x_n is even and let $q = x_n/2$. In this case, it is the numbers

$$-q + 1, -q + 2, \dots, -1, 0, 1, \dots, q - 1, q$$

which are a complete set of mutually incongruent residues modulo x_n . (For, $-q \equiv q \pmod{x_n}$.) As above, $x_{n-1} \leq q$. So the result will follow as above, unless $x_{n-1} = q = x_n/2$, so that $x_{n+1} \equiv -q \pmod{x_n}$, in which case $i = n - 1, j = n + 1$ would contradict our result. But, by Lemma 2.6,

$$x_n = ax_{n-1} + dy_{n-1},$$

so that $x_n = 2x_{n-1}$ implies $a = 2$ and $y_{n-1} = 0$, i.e., $n = 1$. So the result can fail only for $a = 2, n = 1$ and $i = 0, j = 2$.

LEMMA 2.23. *Let $x_j \equiv x_i \pmod{x_n}$, $n > 0$, $0 < i \leq n$, $0 \leq j < 4n$, then either $j = i$ or $j = 4n - i$.*

Proof. First suppose $j \leq 2n$. Then by Lemma 2.22, $j = i$ unless the exceptional case occurs. Since $i > 0$, this can only happen if $j = 0$. But then

$$i = 2 > 1 = n.$$

Otherwise, let $j > 2n$ and set $\bar{j} = 4n - j$ so $0 < \bar{j} < 2n$. By Lemma 2.21, $x_j \equiv x_{\bar{j}} \pmod{x_n}$. Again $\bar{j} = i$ unless the exceptional case of Lemma 2.22 occurs. But this last is out of the question because $i, \bar{j} > 0$.

LEMMA 2.24. *If $0 < i \leq n$ and $x_j \equiv x_i \pmod{x_n}$, then $j \equiv \pm i \pmod{4n}$.*

Proof. Write $j = 4nq + \bar{j}$, $0 \leq \bar{j} < 4n$. By Lemma 2.21,

$$x_i \equiv x_{\bar{j}} \pmod{x_n}.$$

By Lemma 2.23 $i = \bar{j}$ or $i = 4n - \bar{j}$. So, $j \equiv \bar{j} \equiv \pm i \pmod{4n}$.

3. The exponential function. Consider the system of Diophantine equations:

- | | |
|-------|--------------------------|
| (I) | $x^2 - (a^2 - 1)y^2 = 1$ |
| (II) | $u^2 - (a^2 - 1)v^2 = 1$ |
| (III) | $s^2 - (b^2 - 1)t^2 = 1$ |
| (IV) | $v = ry^2$ |
| (V) | $b = 1 + 4py = a + qu$ |
| (VI) | $s = x + cu$ |
| (VII) | $t = k + 4(d - 1)y$ |

$$(VIII) \quad y = k + e - 1.$$

Then it is possible to prove:

THEOREM 3.1. *For given $a, x, k, a > 1$, the system I-VIII has a solution in the remaining arguments $y, u, v, s, t, b, r, p, q, c, d, e$ if and only if $x = x_k(a)$.*

Proof. First let there be given a solution of I-VIII. By V, $b > a > 1$. Then I, II, III imply (by Lemma 2.4) that there are $i, j, n > 0$ such that

$$x = x_i(a), \quad y = y_i(a), \quad u = x_n(a), \quad v = y_n(a), \quad s = x_j(b), \quad t = y_j(b).$$

By IV, $y \leq v$ so that $i \leq n$. V and VI yield the congruences

$$b \equiv a \pmod{x_n(a)}; \quad x_j(b) \equiv x_i(a) \pmod{x_n(a)}$$

and by Lemma 2.15 one gets also

$$x_j(b) \equiv x_j(a) \pmod{x_n(a)}.$$

Thus,

$$x_i(a) \equiv x_j(a) \pmod{x_n(a)}.$$

By Lemma 2.24,

$$(1) \quad j \equiv \pm i \pmod{4n}.$$

Next, equation IV' yields

$$(y_i(a))^2 \mid y_n(a).$$

so that by Lemma 2.12,

$$y_i(a) \mid n$$

and (1) yields:

$$(2) \quad j \equiv \pm i \pmod{4y_i(a)}.$$

By equation V

$$b \equiv 1 \pmod{4y_i(a)},$$

so by Lemma 2.14,

$$(3) \quad y_j(b) \equiv j \pmod{4y_i(a)}.$$

By equation VII,

$$(4) \quad y_j(b) \equiv k \pmod{4y_i(a)}.$$

Combining (2), (3), (4),

$$(5) \quad k \equiv \pm i \pmod{4y_i(a)}.$$

Equation VIII yields

$$k \leq y_i(a)$$

and by Lemma 2.18,

$$i \leq y_i(a).$$

Since the numbers

$$-2y+1, -2y+2, \dots, -1, 0, 1, \dots, 2y$$

form a complete set of mutually incongruent residues modulo $4y = 4y_i(a)$, these inequalities show that (5) implies $k = i$. Hence

$$x = x_i(a) = x_k(a).$$

Conversely, let $x = x_k(a)$. Set $y = y_k(a)$ so that I holds. Let $m = 2ky_k(a)$ and let $u = x_m(a)$, $v = y_m(a)$. Then II is satisfied. By Lemmas 2.9 and 2.11 $y^2 \mid v$. Hence one can choose r satisfying IV. Moreover by Lemma 2.16, v is even so that u is odd. By Lemma 2.7, $(u, v) = 1$. Hence $(u, v, 4y) = 1$. (If p is a prime divisor of u and of $4y$, then $p \mid y$ because u is odd, and hence $p \mid v$ since $y \mid v$.) So by the Chinese Remainder Theorem (Theorem 1.2), one can find b_0 such that

$$b_0 \equiv 1 \pmod{4y}$$

$$b_0 \equiv a \pmod{u}.$$

Since $b_0 + 4juy$ will also satisfy these congruences, b, p, q satisfying V can be found. III is satisfied by setting $s = x_k(b)$, $t = y_k(b)$. Since $b > a$, $s = x_k(b) > x_k(a) = x$. By Lemma 2.15 (using V), $s \equiv x \pmod{u}$. So c can be chosen to satisfy VI. By Lemma 2.18, $t \geq k$ and by Lemma 2.14, $t \equiv k \pmod{b-1}$ and hence using V, $t \equiv k \pmod{4y}$. So d can be chosen to satisfy VII. By Lemma 2.18 again, $y \geq k$, so VIII can be satisfied by setting $e = y - k + 1$.

COROLLARY 3.2. *The function*

$$g(z, k) = x_k(z + 1)$$

is Diophantine.

Proof. Adjoin to the system I-VIII:

$$(A) \quad a = z + 1.$$

By the theorem, the system (A), I-VIII has a solution if and only if $x = x_k(a) = g(z, k)$. Thus a Diophantine definition of g can be obtained in the usual way by summing the squares of 9 polynomials.

Now at last it is possible to prove:

THEOREM 3.3. *The exponential function $h(n, k) = n^k$ is Diophantine.*

First, a simple inequality:

LEMMA 3.4. *If $a > y^k$, then $2ay - y^2 - 1 > y^k$.*

Proof. Set $g(y) = 2ay - y^2 - 1$. Then (since $a \geq 2$) $g(1) = 2a - 2 \geq a$. For $1 \leq y < a$, $g'(y) = 2a - 2y > 0$. So $g(y) \geq a$ for $1 \leq y < a$. Then for $a > y^k \geq y$, $2ay - y^2 - 1 \geq a > y^k$.

Now, adjoin to equations I-VIII:

$$\text{IX} \quad (x - y(a - n) - m)^2 = (f - 1)^2(2an - n^2 - 1)^2$$

$$\text{X} \quad m + g = 2an - n^2 - 1$$

$$\text{XI} \quad w = n + h = k + l$$

$$\text{XII} \quad a^2 - (w^2 - 1)(w - 1)^2 z^2 = 1.$$

Theorem 3.3 then follows at once from:

LEMMA 3.5. *$m = n^k$ if and only if equations I-XII have a solution in the remaining arguments.*

Proof. Suppose I-XII hold. By XI, $w > 1$. Hence $(w - 1)z > 0$ and so by XII $a > 1$. So Theorem 3.1 applies and it follows that $x = x_k(a)$, $y = y_k(a)$. By IX and Lemma 2.17,

$$m \equiv n^k \pmod{2an - n^2 - 1}.$$

XI yields

$$k, n < w.$$

By XII (using Lemma 2.4), for some j , $a = x_j(w)$, $(w - 1)z = y_j(w)$. By Lemma 2.14,

$$j \equiv 0 \pmod{w - 1}$$

so that $j \geq w - 1$. So by Lemma 2.19,

$$a \geq w^{w-1} > n^k.$$

Now by X, $m < 2an - n^2 - 1$, and by Lemma 3.4,

$$n^k < 2an - n^2 - 1.$$

Since m and n^k are congruent and both less than the modulus, they must be equal.

Conversely, suppose that $m = n^k$. Solutions must be found for I-XII. Choose any number w such that $w > n$ and $w > k$. Set $a = x_{w-1}(w)$ so that $a > 1$. By Lemma 2.14,

$$y_{w-1}(w) \equiv 0 \pmod{w - 1}.$$

So one can write

$$y_{w-1}(w) = z(w-1);$$

thus XII is satisfied. XI can be satisfied by setting

$$h = w - n, \quad l = w - k.$$

As before, $a > n^k$ so that again by Lemma 3.4,

$$m = n^k < 2an - n^2 - 1$$

and X can be satisfied. Setting $x = x_k(a)$, $y = y_k(a)$, Lemma 2.17 permits one to define f such that

$$x - y(a - n) - m = \pm (f - 1)(2an - n^2 - 1),$$

so that IX is satisfied. Finally, I-VIII can be satisfied by Theorem 3.1.

4. The language of Diophantine predicates. Now that it has been proved that the *exponential function* is Diophantine, many other functions and sets can be handled. As an example, let

$$h(u, v, w) = u^{v^w}.$$

The claim is that h is a Diophantine function. For:

$$y = u^{v^w} \Leftrightarrow (\exists z) (y = u^z \& z = v^w),$$

where “&” is the logician’s symbol for “and”. Using Theorem 3.3, there is a polynomial P such that:

$$y = u^z \Leftrightarrow (\exists r_1, \dots, r_n) [P(y, u, z, r_1, \dots, r_n) = 0],$$

$$z = v^w \Leftrightarrow (\exists s_1, \dots, s_n) [P(z, v, w, s_1, \dots, s_n) = 0].$$

Then,

$$\begin{aligned} y = u^{v^w} \Leftrightarrow (\exists z, r_1, \dots, r_n, s_1, \dots, s_n) [P^2(y, u, z, r_1, \dots, r_n) \\ + P^2(z, v, w, s_1, \dots, s_n) = 0]. \end{aligned}$$

Now this procedure is perfectly general: Expressions which are already known to yield Diophantine sets may be combined freely using the logical operations of “&” and “(∃)”; the resulting expression will again define a Diophantine set. (Such expressions are sometimes called *Diophantine predicates*.) In this “language” it is also permissible to use the logician’s “∨” for “or”, since:

$$\begin{aligned} (\exists r_1, \dots, r_n) [P_1 = 0] \vee (\exists s_1, \dots, s_m) [P_2 = 0] \\ \Leftrightarrow (\exists r_1, \dots, r_n, s_1, \dots, s_m) [P_1 P_2 = 0]. \end{aligned}$$

Three important Diophantine functions are given by:

THEOREM 4.1. *The following functions are Diophantine:*

- (1) $f(n, k) = \binom{n}{k}$
- (2) $g(n) = n!$
- (3) $h(a, b, y) = \prod_{k=1}^y (a + bk).$

In proving this theorem the familiar notation $[\alpha]$, where α is a real number, will be used to mean the unique integer such that

$$[\alpha] \leq \alpha < [\alpha] + 1.$$

LEMMA 4.1. *For $0 < k \leq n$, $u > 2^n$*

$$[(u+1)^n/u^k] = \sum_{i=k}^n \binom{n}{i} u^{i-k}.$$

Proof.

$$(u+1)^n/u^k = \sum_{i=0}^n \binom{n}{i} u^{i-k} = S + R$$

where

$$S = \sum_{i=k}^n \binom{n}{i} u^{i-k} \quad R = \sum_{i=0}^{k-1} \binom{n}{i} u^{i-k}.$$

Then S is an integer and

$$\begin{aligned} R &< u^{-1} \sum_{i=0}^{k-1} \binom{n}{i} \\ &< u^{-1} \sum_{i=0}^n \binom{n}{i} \\ &= u^{-1}(1+1)^n \\ &< 1. \end{aligned}$$

So,

$$S \leq (u+1)^n/u^k < S + 1$$

which gives the result.

LEMMA 4.2. For $0 < k \leq n$, $u > 2^n$,

$$[(u+1)^n/u^k] \equiv \binom{n}{k} \pmod{u}.$$

Proof. In Lemma 4.1 all terms of the sum for which $i > k$ are divisible by u .

LEMMA 4.3. $f(n, k) = \binom{n}{k}$ is Diophantine.

Proof. Since

$$\binom{n}{k} \leq \sum_{i=0}^n \binom{n}{i} = 2^n < u,$$

Lemma 4.2 determines $\binom{n}{k}$ as the unique positive integer congruent to $[(u+1)^n/u^k]$ modulo u and $< u$. Thus,

$$z = \binom{n}{k} \Leftrightarrow (\exists u, v, w) (v = 2^n \ \& \ u > v \\ \& \ w = [(u+1)^n/u^k] \ \& \ z \equiv w \pmod{u} \ \& \ z < u).$$

To see that $\binom{n}{k}$ is Diophantine, it then suffices to note that each of the above expressions separated by “&” are Diophantine predicates; $v = 2^n$ is of course Diophantine by Theorem 3. The inequality $u > v$ is of course Diophantine since $u > v \Leftrightarrow (\exists x)(u = v + x)$. Also,

$$z \equiv w \pmod{u} \ \& \ z < u \Leftrightarrow (\exists x, y) (w = z + (x-1)u \ \& \ u = z + y).$$

Finally

$$w = [(u+1)^n/u^k]$$

$$\Leftrightarrow$$

$$(\exists x, y, t) (t = u+1 \ \& \ x = t^n \ \& \ y = u^k \ \& \ w \leq x/y < w+1),$$

and $w \leq x/y < w+1 \Leftrightarrow wy \leq x < (w+1)y$.

LEMMA 4.4. If $r > (2x)^{x+1}$ then

$$x! = \left\lfloor r^x / \binom{r}{x} \right\rfloor.$$

Proof. Let $r > (2x)^{x+1}$. Then,

$$r^x / \binom{r}{x} = \frac{r^x x!}{r(r-1) \cdots (r-x+1)} \\ = x! \left\{ \frac{1}{\left(1 - \frac{1}{r}\right) \cdots \left(1 - \frac{x-1}{r}\right)} \right\}$$

$$< x! \cdot \frac{1}{\left(1 - \frac{x}{r}\right)^x}.$$

Now,

$$\begin{aligned} \frac{1}{1 - \frac{x}{r}} &= 1 + \frac{x}{r} + \left(\frac{x}{r}\right)^2 + \dots \\ &= 1 + \frac{x}{r} \left\{ 1 + \frac{x}{r} + \left(\frac{x}{r}\right)^2 + \dots \right\} \\ &< 1 + \frac{x}{r} \left\{ 1 + \frac{1}{2} + \frac{1}{4} + \dots \right\} \\ &= 1 + \frac{2x}{r}. \end{aligned}$$

And,

$$\begin{aligned} \left(1 + \frac{2x}{r}\right)^x &= \sum_{j=0}^x \binom{x}{j} \left(\frac{2x}{r}\right)^j \\ &< 1 + \frac{2x}{r} \sum_{j=1}^x \binom{x}{j} \\ &< 1 + \frac{2x}{r} \cdot 2^x. \end{aligned}$$

So,

$$\begin{aligned} r^x / \binom{r}{x} &< x! + \frac{2x}{r} \cdot x! 2^x \\ &< x! + \frac{2^{x+1} x^{x+1}}{r} \\ &< x! + 1. \end{aligned}$$

LEMMA 4.5. $n!$ is a Diophantine function.

Proof. $m = n! \Leftrightarrow$

$$\begin{aligned} (\exists r, s, t, u, v) \{ &s = 2x + 1 \text{ \& } t = x + 1 \text{ \& } r = s^t \\ &\& u = r^n \text{ \& } v = \binom{r}{n} \text{ \& } mv \leq u < (m + 1)v \}. \end{aligned}$$

LEMMA 4.6. Let $bq \equiv a \pmod{M}$. Then,

$$\prod_{k=1}^y (a + bk) \equiv b^y y! \binom{q + y}{y} \pmod{M}.$$

Proof.

$$\begin{aligned}
 b^y y! \binom{q+y}{y} &= b^y (q+y) (q+y-1) \cdots (q+1) \\
 &= (bq + yb) (bq + (y-1)b) \cdots (bq + b) \\
 &\equiv (a + yb) (a + (y-1)b) \cdots (a + b) \pmod{M}.
 \end{aligned}$$

LEMMA 4.7. $h(a, b, y) = \prod_{k=1}^y (a + bk)$ is a Diophantine function.

Proof. In Lemma 4.6 choose $M = b(a + by)^y + 1$. Then, $(M, b) = 1$ and $M > \prod_{k=1}^y (a + bk)$. Hence the congruence $bq \equiv a \pmod{M}$ is solvable for q and then $\prod_{k=1}^y (a + bk)$ is determined as the unique number which is congruent modulo M to $b^y y! \binom{q+y}{y}$ and is also $< M$. I.e.,

$$\begin{aligned}
 z = \prod_{k=1}^y (a + bk) &\Leftrightarrow (\exists M, p, q, r, s, t, u, v, w, x) \\
 &\left\{ \begin{aligned} &r = a + by \ \& \ s = r^y \ \& \ M = bs + 1 \\ &\& \ bq = a + Mt \ \& \ u = b^y \ \& \ v = y! \ \& \ z < M \\ &\& \ w = q + y \ \& \ x = \binom{w}{y} \ \& \ z + Mp = uvx \end{aligned} \right\}.
 \end{aligned}$$

Using the previous expressions for the exponential function, for $v = y!$ and for $x = \binom{w}{y}$, we obtain the result.

The assertion of Theorem 4.1 is contained in Lemmas 4.3, 4.5, and 4.7.

5. Bounded quantifiers. The language of Diophantine predicates permits use of $\&$, \vee , and \exists . Other operations used by logicians are:

$$\begin{aligned}
 \sim &\quad \text{for "not"} \\
 (\forall x) &\quad \text{for "for all } x\text{"} \\
 \rightarrow &\quad \text{for "if } \dots, \text{ then } \dots\text{"}
 \end{aligned}$$

However, as will be clear later, the use of any of these other operations can lead to expressions which define sets that are not Diophantine. There are also the *bounded existential quantifiers*:

$$"(\exists y)_{\leq x} \dots" \text{ which means } "(\exists y) (y \leq x \ \& \ \dots)"$$

and the *bounded universal quantifiers*:

$$"(\forall y)_{\leq x} \dots" \text{ which means } "(\forall y) (y > x \vee \dots)".$$

It turns out that these operations may be adjoined to the language of Diophantine predicates; that is, the sets defined by expressions of this extended language will still be Diophantine. I.e.,

THEOREM 5.1. *If P is a polynomial,*

$$R = \{ \langle y, x_1, \dots, x_n \rangle \mid (\exists z)_{\leq y} (\exists y_1, \dots, y_m) [P(y, z, x_1, \dots, x_n, y_1, \dots, y_m) = 0] \}$$

and

$$S = \{ \langle y, x_1, \dots, x_n \rangle \mid (\forall z)_{\leq y} (\exists y_1, \dots, y_m) [P(y, z, x_1, \dots, x_n, y_1, \dots, y_m) = 0] \},$$

then R and S are Diophantine.

That R is Diophantine is trivial. Namely,

$$\langle y, x_1, \dots, x_n \rangle \in R \Leftrightarrow (\exists z, y_1, \dots, y_m) (z \leq y \ \& \ P = 0).$$

The proof of the other half of the theorem is far more complicated.

LEMMA 5.1.

$$(\forall k)_{\leq y} (\exists y_1, \dots, y_m) [P(y, k, x_1, \dots, x_n, y_1, \dots, y_m) = 0]$$

$$\Leftrightarrow$$

$$(\exists u) (\forall k)_{\leq y} (\exists y_1, \dots, y_m)_{\leq u} [P(y, k, x_1, \dots, x_n, y_1, \dots, y_m) = 0].$$

Proof. The right side of the equivalence trivially implies the left side. For the converse, suppose the left side is true for given y, x_1, \dots, x_n . Then for each $k = 1, 2, \dots, y$ there are definite numbers $y_1^{(k)}, \dots, y_m^{(k)}$ for which:

$$P(y, k, x_1, \dots, x_n, y_1^{(k)}, \dots, y_m^{(k)}) = 0.$$

Taking u to be the maximum of the my numbers

$$\{y_j^{(k)} \mid j = 1, \dots, m; k = 1, 2, \dots, y\},$$

it follows that the right side of the equivalence is likewise true.

LEMMA 5.2. *Let $Q(y, u, x_1, \dots, x_n)$ be a polynomial with the properties:*

- (1) $Q(y, u, x_1, \dots, x_n) > u$, (2) $Q(y, u, x_1, \dots, x_n) > y$,
 (3) $k \leq y$ and $y_1, \dots, y_m \leq u$ imply $|P(y, k, x_1, \dots, x_n, y_1, \dots, y_m)| \leq Q(y, u, x_1, \dots, x_n)$.
 Then,

$$(\forall k)_{\leq y} (\exists y_1, \dots, y_m)_{\leq u} [P(y, k, x_1, \dots, x_n, y_1, \dots, y_m) = 0]$$

$$\Leftrightarrow$$

$$(\exists c, t, a_1, \dots, a_m) [1 + ct = \prod_{k=1}^y (1 + kt)]$$

$$\& t = Q(y, u, x_1, \dots, x_n)! \& 1 + ct \mid \prod_{j=1}^u (a_1 - j)$$

$$\& \dots \& 1 + ct \mid \prod_{j=1}^u (a_m - j)$$

$$\& P(y, c, x_1, \dots, x_n, a_1, \dots, a_m) \equiv 0 \pmod{1 + ct}.$$

The point of this lemma is that while the right side of the equivalence seems the more complicated of the two, it is free of bounded universal quantifiers.

Proof. First the implication in the \Leftarrow direction:

For each $k = 1, 2, \dots, y$, let p_k be a prime factor of $1 + kt$. Let $y_i^{(k)}$ be the remainder when a_i is divided by p_k ($k = 1, 2, \dots, y$; $i = 1, 2, \dots, m$). It will follow that for each k, i :

$$(a) \quad 1 \leq y_i^{(k)} \leq u$$

$$(b) \quad P(y, k, x_1, \dots, x_n, y_1^{(k)}, \dots, y_m^{(k)}) = 0.$$

To demonstrate (a), note that $p_k \mid 1 + kt, 1 + kt \mid 1 + ct$ and $1 + ct \mid \prod_{j=1}^u (a_i - j)$. I.e., $p_k \mid \prod_{j=1}^u (a_i - j)$. Since p_k is a prime, $p_k \mid a_i - j$ for some $j = 1, 2, \dots, u$. That is

$$j \equiv a_i \equiv y_i^{(k)} \pmod{p_k}.$$

Since $t = Q(y, u, x_1, \dots, x_n)!$, (2) implies that every divisor of $1 + kt$ must be $> Q(y, u, x_1, \dots, x_n)$. So $p_k > Q(y, u, x_1, \dots, x_n)$ and by (1), $p_k > u$. Hence $j \leq u < p_k$. Since $y_i^{(k)}$ is the remainder when a_i is divided by p_k , also $y_i^{(k)} < p_k$. So,

$$y_i^{(k)} = j.$$

To demonstrate (b), first note that

$$1 + ct \equiv 1 + kt \equiv 0 \pmod{p_k}.$$

Hence

$$k + kct \equiv c + kct \pmod{p_k},$$

i.e., $k \equiv c \pmod{p_k}$. We have already obtained

$$y_i^{(k)} \equiv a_i \pmod{p_k}.$$

Thus,

$$\begin{aligned} P(y, k, x_1, \dots, x_n, y_1^{(k)}, \dots, y_m^{(k)}) &\equiv P(y, c, x_1, \dots, x_n, a_1, \dots, a_m) \\ &\equiv 0 \pmod{p_k}. \end{aligned}$$

Finally

$$|P(y, k, x_1, \dots, x_n, y_1^{(k)}, \dots, y_m^{(k)})| \leq Q(y, u, x_1, \dots, x_n) < p_k.$$

This proves (b) and completes the proof of the \Leftarrow implication.

To prove the \Rightarrow implication, let

$$P(y, k, x_1, \dots, x_n, y_1^{(k)}, \dots, y_m^{(k)}) = 0,$$

for each $k = 1, 2, \dots, t$, where each $y_j^{(k)} \leq u$. We set $t = Q(y, u, x_1, \dots, x_n)!$, and since $\prod_{k=1}^y (1 + kt) \equiv 1 \pmod t$, we can find c such that

$$1 + ct = \prod_{k=1}^y (1 + kt).$$

Now, it is claimed that for $1 \leq k < l \leq y$,

$$(1 + kt, 1 + lt) = 1.$$

For, let $p \mid 1 + kt, p \mid 1 + lt$. Then $p \mid l - k$, so $p < y$. But since $Q(y, u, x_1, \dots, x_n) > y$ this implies $p \mid t$ which is impossible. Thus the numbers $1 + kt$ form an *admissible sequence of moduli* and the Chinese Remainder Theorem (Theorem 1.2) may be applied to yield, for each i , $1 \leq i \leq m$, a number a_i such that

$$a_i \equiv y_i^{(k)} \pmod{1 + kt}, \quad k = 1, 2, \dots, y.$$

As above, $k \equiv c \pmod{1 + kt}$. So

$$\begin{aligned} P(y, c, x_1, \dots, x_n, a_1, \dots, a_m) &\equiv P(y, k, x_1, \dots, x_n, y_1^{(k)}, \dots, y_m^{(k)}) \pmod{1 + kt}, \\ &= 0. \end{aligned}$$

Since the numbers $1 + kt$ are relatively prime in pairs and each divides $P(y, c, x_1, \dots, x_n, a_1, \dots, a_m)$ so does their product. I.e.,

$$P(y, c, x_1, \dots, x_n, a_1, \dots, a_m) \equiv 0 \pmod{1 + ct}.$$

Finally,

$$a_i \equiv y_i^{(k)} \pmod{1 + kt},$$

i.e.,

$$1 + kt \mid a_i - y_i^{(k)}.$$

Since $1 \leq y_i^{(k)} \leq u$,

$$1 + kt \mid \prod_{j=1}^u (a_i - j).$$

And again since the $1 + kt$'s are relatively prime to one another,

$$1 + ct \mid \prod_{j=1}^u (a_i - j).$$

Now it is easy to complete the proof of Theorem 5.1 using Lemmas 5.1 and 5.2. First find a polynomial Q satisfying (1), (2), (3) of Lemma 5.2. This is easy to do: Write

$$P(y, k, x_1, \dots, x_n, y_1, \dots, y_m) = \sum_{r=1}^N t_r$$

where each t_r has the form

$$t_r = |c| y^a k^b x_1^{q_1} x_2^{q_2} \dots x_n^{q_n} y_1^{s_1} y_2^{s_2} \dots y_m^{s_m}$$

for c an integer positive or negative. Set $u_r = c y^{a+b} x_1^{q_1} x_2^{q_2} \dots x_n^{q_n} u^{s_1+s_2+\dots+s_m}$ and let

$$Q(y, u, x_1, \dots, x_n) = u + y + \sum_{r=1}^N u_r.$$

Then (1), (2), and (3) of Lemma 5.2 hold trivially. Thus:

$$(\forall k)_{\leq y} (\exists y_1, \dots, y_m) [P(y, k, x_1, \dots, x_n, y_1, \dots, y_m) = 0]$$

$$\Leftrightarrow$$

$$(\exists u, c, t, a_1, \dots, a_m) \left[1 + ct = \prod_{k=1}^y (1 + kt) \right.$$

$$\& t = Q(y, u, x_1, \dots, x_n)! \& 1 + ct \mid \prod_{j=1}^u (a_1 - j)$$

$$\& \dots \& 1 + ct \mid \prod_{j=1}^u (a_m - j)$$

$$\& P(y, c, x_1, \dots, x_n, a_1, \dots, a_m) \equiv 0 \pmod{1 + ct} \Big]$$

$$\Leftrightarrow$$

$$(\exists u, c, t, a_1, \dots, a_n, e, f, g_1, \dots, g_m, h_1, \dots, h_n, l)$$

$$\left[e = 1 + ct \& e = \prod_{k=1}^y (1 + kt) \& f = Q(y, u, x_1, \dots, x_n) \right.$$

$$\& t = f! \& g_1 = a_1 - u - 1 \& g_2 = a_2 - u - 1 \& \dots \& g_m = a_m - u - 1$$

$$\& h_1 = \prod_{k=1}^u (g_1 + k) \& h_2 = \prod_{k=1}^u (g_2 + k)$$

$$\& \dots \& h_m = \prod_{k=1}^u (g_m + k) \& e \mid h_1 \& e \mid h_2 \& \dots \& e \mid h_m$$

$$\& l = P(y, c, x_1, \dots, x_n, a_1, \dots, a_n) \& e \mid l \Big]$$

and this is Diophantine by Theorem 4.1.

6. Recursive functions. So far one trick after another has been used to show that various sets are Diophantine. But now very powerful methods are available: it turns out that the expanded version of the language of Diophantine predicates, permitting the use of bounded quantifiers (sanctioned by Theorem 5.1) together with the Sequence Number Theorem (Theorem 1.3) enables one to show in quite a straightforward way that almost any set we please is Diophantine.

Some examples are in order:

(i) *the set P of prime numbers:*

$$x \in P \Leftrightarrow x > 1 \ \& \ (\forall y, z)_{\leq x} [yz < x \vee yz > x \vee y = 1 \vee z = 1].$$

Another Diophantine definition of the primes is:

$$\begin{aligned} x \in P &\Leftrightarrow x > 1 \ \& \ ((x-1)!, x) = 1 \\ &\Leftrightarrow x > 1 \ \& \ (\exists y, z, u, v) [y = x-1 \ \& \ z = y! \ \& \ (uz - vx)^2 = 1]; \end{aligned}$$

but the first definition is the more natural one.

From Theorem 1.4 it follows that *there is a "prime-representing" polynomial P* , i.e., a positive integer is prime if and only if it is in the range of P . For an explicit construction of such a polynomial P , cf. [23a].

(ii) *the function $g(y) = \prod_{k=1}^y (1 + k^2)$.* Here we use the Sequence Number Theorem to "encode" the sequence $g(1), g(2), \dots, g(y)$ into a single number u , i.e., so that

$$S(i, u) = g(i), \quad i = 1, 2, \dots, y.$$

Thus, $z = g(y)$

$$\begin{aligned} &\Leftrightarrow (\exists u) \{S(1, u) = 2 \ \& \ (\forall k)_{\leq y} [k = 1 \vee (S(k, u) = (1 + k^2)S(k-1, u))] \ \& \ z = S(y, u)\} \\ &\Leftrightarrow (\exists u) \{S(1, u) = 2 \ \& \ (\forall k)_{\leq y} [k = 1 \vee (\exists a, b, c) (a = k-1 \\ &\quad \& \ b = S(a, u) \ \& \ c = S(k, u) \ \& \ c = (1 + k^2)b)] \ \& \ z = S(y, u)\}. \end{aligned}$$

By now it is clear that the available methods are quite general. They are so powerful that the question becomes: how can any "reasonable" set or function escape these methods, i.e., not be Diophantine?

The strength of the methods can be tested by considering the class of all *computable* or *recursive* functions. These are the functions which can be computed by a finite program or computing machine having arbitrarily large amounts of time and memory at its disposal. Many rigorous definitions of this class (all of them equivalent) are available. One of the simplest is as follows:

The *recursive functions*³ are all those functions obtainable from the initial functions

$$c(x) = 1, \quad s(x) = x + 1; \quad U_i^n(x_1, \dots, x_n) = x_i, \quad 1 \leq i \leq n;$$

$$S(i, u) \text{ (The sequence number function)}^4$$

iteratively applying the three operations: *composition*, *primitive recursion*, and *minimalization* defined below:

COMPOSITION yields the function

$$h(x_1, \dots, x_n) = f(g_1(x_1, \dots, x_n), \dots, g_m(x_1, \dots, x_n))$$

from the given functions g_1, \dots, g_m and $f(t_1, \dots, t_m)$.

PRIMITIVE RECURSION yields the function $h(x_1, \dots, x_n, z)$ which satisfies the equations:

$$h(x_1, \dots, x_n, 1) = f(x_1, \dots, x_n)$$

$$h(x_1, \dots, x_n, t + 1) = g(t, h(x_1, \dots, x_n, t), x_1, \dots, x_n),$$

from the given functions f, g .

When $n = 0$, f becomes a constant so that h is obtained directly from g .

MINIMALIZATION yields the function:

$$h(x_1, \dots, x_n) = \min_y [f(x_1, \dots, x_n, y) = g(x_1, \dots, x_n, y)]$$

from the given functions f, g assuming that f, g are such that for each x_1, \dots, x_n there is at least one y satisfying the equation $f(x_1, \dots, x_n, y) = g(x_1, \dots, x_n, y)$; (i.e., h must be everywhere defined).

The main result of this article is:

THEOREM 6.1. *A function is Diophantine if and only if it is recursive.*

To begin with, consider the following short list of recursive functions:

(1) $x + y$ is recursive since

$$x + 1 = s(x),$$

$$x + (t + 1) = s(x + t) = g(t, x + t, x),$$

where $g(u, v, w) = s(U_2^3(u, v, w))$.

(2) $x \cdot y$ is recursive since

$$x \cdot 1 = U_1^1(x)$$

$$x \cdot (t + 1) = (x \cdot t) + x = g(t, x \cdot t, x),$$

where $g(u, v, w) = U_2^3(u, v, w) + U_3^3(u, v, w)$.

(3) For each fixed k , the constant function $c_k(x) = k$ is recursive, since $c_1(x)$ is one of the initial functions and $c_{k+1}(x) = c_k(x) + c(x)$.

(4) Any polynomial $P(x_1, \dots, x_n)$ with *positive* integer coefficients is recursive, since any such function can be expressed by a finite iteration of additions and multiplications of variables and $c(x)$. E.g.,

$$2x^2y + 3xz^3 + 5 = c_2(x) \cdot x \cdot x \cdot y + c_3(x) \cdot x \cdot z \cdot z \cdot z + c_5(x).$$

So (1), (2), (3), and composition gives the result.

Now it is easy to see that every Diophantine function is recursive:

Let f be Diophantine, and write:

$$\begin{aligned} y = f(x_1, \dots, x_n) &\Leftrightarrow (\exists t_1, \dots, t_m) [P(x_1, \dots, x_n, y, t_1, \dots, t_m) \\ &= Q(x_1, \dots, x_n, y, t_1, \dots, t_m)], \end{aligned}$$

where P, Q are polynomials with *positive* integer coefficients. Then, by the sequence number theorem:

$$\begin{aligned} f(x_1, \dots, x_n) &= S(1, \min_u [P(x_1, \dots, x_n, S(1, u), S(2, u), \dots, S(m+1, u)) \\ &= Q(x_1, \dots, x_n, S(1, u), S(2, u), \dots, S(m+1, u))]). \end{aligned}$$

Since $P, Q, S(i, u)$ are recursive, so is f (using composition and minimalization).

To obtain the converse: $S(i, u)$ is known to be Diophantine; the other initial functions are trivially Diophantine. Hence it suffices to prove that the Diophantine functions are closed under composition, primitive recursion and minimalization.

Composition: If $h(x_1, \dots, x_n) = f(g_1(x_1, \dots, x_n), \dots, g_m(x_1, \dots, x_n))$, where f, g_1, \dots, g_m are Diophantine, then so is h since

$$\begin{aligned} y = h(x_1, \dots, x_n) &\Leftrightarrow (\exists t_1, \dots, t_m) [t_1 = g_1(x_1, \dots, x_n) \& \dots \\ &\& t_m = g_m(x_1, \dots, x_n) \& y = f(t_1, \dots, t_m)]. \end{aligned}$$

Primitive Recursion: If

$$\begin{aligned} h(x_1, \dots, x_n, 1) &= f(x_1, \dots, x_n) \\ h(x_1, \dots, x_n, t+1) &= g(t, h(x_1, \dots, x_n, t), x_1, \dots, x_n), \end{aligned}$$

and f, g are Diophantine, then (using the sequence number theorem to "code" the numbers $h(x_1, \dots, x_n, 1), h(x_1, \dots, x_n, 2), \dots, h(x_1, \dots, x_n, z)$):

$$\begin{aligned} y = h(x_1, \dots, x_n, z) &\Leftrightarrow \\ (\exists u) \{ (\exists v) [v = S(1, u) \& v = f(x_1, \dots, x_n)] \\ &\& (\forall t)_{\leq z} [(t = z) \vee (\exists v) (v = S(t+1, u) \\ &\& v = g(t, S(t, u), x_1, \dots, x_n))] \& y = S(z, u) \} \end{aligned}$$

so that (using Theorem 5.1) h is Diophantine.

Minimalization: If

$$h(x_1, \dots, x_n) = \min_y [f(x_1, \dots, x_n, y) = g(x_1, \dots, x_n, y)],$$

where f, g are Diophantine, then so is h since,

$$\begin{aligned} y = h(x_1, \dots, x_n) &\Leftrightarrow \\ (\exists z) [z = f(x_1, \dots, x_n, y) \ \& \ z = g(x_1, \dots, x_n, y)] \\ \& \ (\forall t)_{\leq y} [(t = y) \vee (\exists u, v) (u = f(x_1, \dots, x_n, t) \\ \& \ v = g(x_1, \dots, x_n, t) \ \& \ (u < v \vee v < u))]. \end{aligned}$$

7. A universal Diophantine set. An explicit enumeration of all the Diophantine sets of positive integers will now be described. Any polynomial with positive integer coefficients can be built up from 1 and variables by successive additions and multiplications. We fix the alphabet

$$x_0, x_1, x_2, x_3, \dots$$

of variables and then set up the following enumeration of all such polynomials (using the pairing functions):

$$\begin{aligned} P_1 &= 1 \\ P_{3i-1} &= x_{i-1} \\ P_{3i} &= P_{L(i)} + P_{R(i)} \\ P_{3i+1} &= P_{L(i)} \cdot P_{R(i)}. \end{aligned}$$

Write $P_i = P_i(x_0, x_1, \dots, x_n)$, where n is large enough so that all variables occurring in P_i are included. (Of course P_i will not in general depend on all of these variables.) Finally, let

$$D_n = \{x_0 \mid (\exists x_1, \dots, x_n) [P_{L(n)}(x_0, x_1, \dots, x_n) = P_{R(n)}(x_0, x_1, \dots, x_n)]\}.$$

Here, $P_{L(n)}$ and $P_{R(n)}$ do not actually involve all of the variables x_0, x_1, \dots, x_n —but clearly cannot involve any others. (Recall that $L(n), R(n) \leq n$.) By the way the sequence P_i has been constructed, it is seen that the sequence of sets:

$$D_1, D_2, D_3, D_4, \dots$$

includes all Diophantine sets. Moreover:

THEOREM 7.1 (Universality Theorem⁵).

$$\{\langle n, x \rangle \mid x \in D_n\} \text{ is Diophantine.}$$

Proof. Once again using the sequence number theorem, it is claimed that:

$$\begin{aligned} x \in D_n &\Leftrightarrow (\exists u) \{S(1, u) = 1 \ \& \ S(2, u) = x \\ &\& (\forall i)_{\leq n} [S(3i, u) = S(L(i), u) + S(R(i), u)] \\ &\& (\forall i)_{\leq n} [S(3i + 1, u) = S(L(i), u) \cdot S(R(i), u)] \\ &\& S(L(n), u) = S(R(n), u)\}. \end{aligned}$$

It is clear enough that the predicate on the right-hand side of this equivalence is Diophantine, so it is only necessary to verify the claim:

Let $x \in D_n$ for given x, n . Then there are numbers t_1, \dots, t_n such that $P_{L(n)}(x, t_1, \dots, t_n) = Q_{L(n)}(x, t_1, \dots, t_n)$. Choose u (by the sequence number theorem) so that

$$(*) \quad S(j, u) = P_j(x, t_1, \dots, t_n), \quad j = 1, 2, \dots, 3n + 2.$$

Then in particular $S(2, u) = x$ and $S(3i - 1, u) = t_{i-1}$, $i = 2, 3, \dots, n + 1$. Thus the right-hand side of the equivalence is true.

Conversely, let the right-hand side hold for given n, x . Set

$$t_1 = S(5, u), \ t_2 = S(8, u), \ \dots, \ t_n = S(3n + 2, u).$$

Then, (*) must be true. Since $S(L(n), u) = S(R(n), u)$, it must be the case that

$$P_{L(n)}(x, t_1, \dots, t_n) = P_{R(n)}(x, t_1, \dots, t_n),$$

so that $x \in D_n$.

Since D_1, D_2, D_3, \dots , gives an enumeration of all Diophantine sets, it is easy to construct a set different from all of them and hence non-Diophantine. That is, define:

$$V = \{n \mid n \notin D_n\}.$$

THEOREM 7.2. *V is not Diophantine.*

Proof. This is a simple application of Cantor's diagonal method. If V were Diophantine, then for some fixed i , $V = D_i$. Does $i \in V$? We have:

$$i \in V \Leftrightarrow i \in D_i; \quad i \in V \Leftrightarrow i \notin D_i.$$

This is a contradiction.

THEOREM 7.3. *The function $g(n, x)$ defined by:*

$$\begin{aligned} g(n, x) &= 1 \quad \text{if } x \notin D_n, \\ g(n, x) &= 2 \quad \text{if } x \in D_n, \end{aligned}$$

is not recursive.

Proof. If g were recursive then it would be Diophantine (Theorem 6.1), say:

$$y = g(n, x) \Leftrightarrow (\exists y_1, \dots, y_m) [P(n, x, y, y_1, \dots, y_m) = 0].$$

But then, it would follow that

$$V = \{x \mid (\exists y_1, \dots, y_m) [P(x, x, 1, y_1, \dots, y_m) = 0]\}$$

which contradicts Theorem 7.2.

Using Theorem 7.1, write:

$$x \in D_n \Leftrightarrow (\exists z_1, \dots, z_k) [P(n, x, z_1, \dots, z_k) = 0].$$

where P is some definite (though complicated) polynomial. Suppose there were an algorithm for testing Diophantine equations for possession of positive integer solutions; i.e., *an algorithm for Hilbert's tenth problem!* Then for given n, x this algorithm could be used to test whether or not the equation

$$P(n, x, z_1, \dots, z_k) = 0$$

has a solution, i.e., whether or not $x \in D_n$. Thus the algorithm could be used to compute the function $g(n, x)$. Since the recursive functions are just those for which a computing algorithm exists, g would have to be recursive. This would contradict Theorem 7.3, and this contradiction proves:

THEOREM 7.4. *Hilbert's tenth problem is unsolvable!*

Naturally this result gives no information about the existence of solutions for any *specific* Diophantine equation; it merely guarantees that there is no single algorithm for testing the class of all Diophantine equations. Also note that:

$$\begin{aligned} x \in V &\Leftrightarrow \sim (\exists z_1, \dots, z_k) [P(x, x, z_1, \dots, z_k) = 0] \\ &\Leftrightarrow \{(\exists z_1, \dots, z_k) [P(x, x, z_1, \dots, z_k) = 0] \rightarrow 1 = 0\} \\ &\Leftrightarrow (\forall z_1, \dots, z_k) [P(x, x, z_1, \dots, z_k) > 0] \\ &\quad \vee P(x, x, z_1, \dots, z_k) < 0] \end{aligned}$$

which shows that if either \sim or unbounded universal quantifiers ($\forall z$) or implication (\rightarrow) are permitted in the language of Diophantine predicates, then non-Diophantine sets will be produced.

It is natural to associate with each Diophantine set a *dimension* and a *degree*; i.e., the *dimension of S* is the least n for which a polynomial P exists for which:

$$(*) \quad S = \{x \mid (\exists y_1, \dots, y_n) [P(x, y_1, \dots, y_n) = 0]\},$$

and the *degree of S* is the least degree of a polynomial P satisfying (*) (permitting n to be as large as one likes). Now it is easy to see:

THEOREM 7.5. *Every Diophantine set has degree ≤ 4 .*

Proof. The degree of P satisfying (*) may be reduced by introducing additional

variables z_j satisfying equations of the form

$$z_j = y_i y_k$$

$$z_j = y_i^2$$

$$z_j = x y_i$$

$$z_j = x^2.$$

By successive substitutions of the z_j 's into P its degree can be brought down to 2. Hence the equation is equivalent to a system of simultaneous equations each of degree 2. Summing the squares gives an equation of degree 4.

A less trivial (and more surprising) fact is:

THEOREM 7.6. *There is an integer m such that every Diophantine set has dimension $\leq m$.*

Proof. Write

$$D_n = \{x \mid (\exists y_1, \dots, y_m) [P(x, n, y_1, \dots, y_m) = 0]\},$$

which is possible by the universality theorem. Then the dimension of D_n is $\leq m$ for all n .

An interesting example is given by the sequence of Diophantine sets:

$$S_q = \{x \mid (\exists y_1, \dots, y_q) [x = (y_1 + 1) \cdots (y_q + 1)]\}.$$

Here S_2 is the set of composite numbers; S_q is the set of “ q -fold” composite numbers. It is surely surprising that it is possible to give a Diophantine definition of S_q (for large q) requiring fewer than q parameters (cf. [19]).

How large is m , the number of parameters in the universal Diophantine set? A direct calculation using the arguments given here would yield a number around 50. Actually Matiyacevič and Julia Robinson have very recently shown that $m = 14$ will suffice!

The unsolvability of Hilbert's tenth problem can be used to obtain a strengthened form of Gödel's famous incompleteness theorem:

THEOREM 7.7. *Corresponding to any given axiomatization of number theory, there is a Diophantine equation which has no positive integer solutions, but such that this fact cannot be proved within the given axiomatization.*

A rigorous proof would involve a precise definition of “axiomatization of number theory” which is outside the scope of this article. An informal heuristic argument follows:

One uses the given axiomatization to systematically generate all of the theorems (i.e., consequences of the axioms). Among these theorems will be some asserting

that some Diophantine equation has no solution. Whenever such is encountered it is placed on a special list called LIST A. At the same time a list, LIST B, is made of Diophantine equations which have solutions. LIST B is constructed by a search procedure, e.g., at the n th stage of the search look at the first n Diophantine equations (in a suitable list) and test for solutions in which each argument is $\leq n$. Thus every Diophantine equation which has positive integer solutions will eventually be placed in LIST B. If likewise each Diophantine equation with no solutions would eventually appear in LIST A, then one would have an algorithm for Hilbert's tenth problem. Namely, to test a given equation for possession of a solution simply begin generating LIST A and LIST B until the given equation appears in one list or the other. Since Hilbert's tenth problem is unsolvable, some equation with no solution must be omitted from LIST A. But this is just the assertion of the theorem.

8. Recursively enumerable sets. It is now time to settle the question raised at the beginning: which sets are Diophantine?

DEFINITION. 8.1. *A set S of n -tuples of positive integers is called recursively enumerable if there are recursive functions $f(x, x_1, \dots, x_n), g(x, x_1, \dots, x_n)$ such that:*

$$S = \{ \langle x_1, \dots, x_n \rangle \mid (\exists x) [f(x, x_1, \dots, x_n) = g(x, x_1, \dots, x_n)] \}.$$

THEOREM 8.1. *A set S is Diophantine if and only if it is recursively enumerable.*

Proof. If S is Diophantine there are polynomials P, Q with positive coefficients such that:

$$\begin{aligned} \langle x_1, \dots, x_n \rangle \in S &\Leftrightarrow (\exists y_1, \dots, y_m) [P(x_1, \dots, x_n, y_1, \dots, y_m) = Q(x_1, \dots, x_n, y_1, \dots, y_m)] \\ &\Leftrightarrow (\exists u) [P(x_1, \dots, x_n, S(1, u), \dots, S(m, u)) = Q(x_1, \dots, x_n, S(1, u), \dots, S(m, u))], \end{aligned}$$

so that S is recursively enumerable.

Conversely if S is recursively enumerable there are recursive functions $f(x, x_1, \dots, x_n), g(x, x_1, \dots, x_n)$ such that

$$\begin{aligned} \langle x_1, \dots, x_n \rangle \in S &\Leftrightarrow (\exists x) [f(x, x_1, \dots, x_n) = g(x, x_1, \dots, x_n)] \\ &\Leftrightarrow (\exists x, z) [z = f(x, x_1, \dots, x_n) \ \& \ z = g(x, x_1, \dots, x_n)]. \end{aligned}$$

Thus by Theorem 6.1, S is Diophantine.

9. Historical appendix. The present exposition has ignored the chronological order in which the ideas were developed. The first contribution was by Gödel in his celebrated 1931 paper [16]. The main point of Gödel's investigation was the existence of undecidable statements in formal systems. The undecidable statements Gödel obtained involved recursive functions, and in order to exhibit the simple number-theoretic character of these statements, Gödel used the Chinese remainder theorem to reduce them to "arithmetic" form. The technique used is just what is

used here in proving Theorem 1.3 (the sequence number theorem) and Theorem 6.1 (in the direction: every recursive function is Diophantine). However without the techniques for dealing with bounded universal quantifiers as discussed in this paper, the best result yielded by Gödel's methods is that every recursive function (and indeed every recursively enumerable set) can be defined by a Diophantine equation preceded by a finite number of existential and bounded universal quantifiers⁶. In my doctoral dissertation (cf. [5], [6]), I showed that all but one of the bounded universal quantifiers could be eliminated, so that every recursively enumerable set S could be defined as

$$S = \{x \mid (\exists y) (\forall k)_{\leq y} (\exists y_1, \dots, y_m) [P(k, x, y, y_1, \dots, y_m) = 0]\}.$$

This representation became known as the Davis normal form. (Later R. M. Robinson [31], [32] showed that in this normal form one could take $m = 4$. More recently Matiyacevič has shown that one can even take $m = 2$. It is known that one cannot always have $m = 0$; whether one can always get $m = 1$ is open.)

Independent of my work and at about the same time, Julia Robinson began her study [27] of Diophantine sets. Her investigations centered about the question: *Is the exponential function Diophantine?* The main result was that a certain hypothesis implied that the exponential function was Diophantine. The hypothesis, which became known as the Julia Robinson hypothesis, has played a key role in work on Hilbert's tenth problem. Its statement is simply:

There exists a Diophantine set D such that:

- (1) $\langle u, v \rangle \in D$ implies $v \leq u^u$.
- (2) For each k , there is $\langle u, v \rangle \in D$ such that $v > u^k$.

The hypothesis remained an open question for about 2 decades. (Actually the set

$$D = \{\langle u, v \rangle \mid v = x_u(2) \text{ \& } u > 3\}$$

satisfies (1) and (2) by Lemma 2.19 and is Diophantine by Corollary 3.2, so the truth of Julia Robinson's hypothesis follows at once from the results in this article.) Julia Robinson's proof that this hypothesis implies that the exponential function is Diophantine used the Pell equation. And, the proof that the exponential function is indeed Diophantine given here is closely related to a more recent proof [28] by her of this same implication.

In [27], Julia Robinson studied also sets and functions which were *exponential Diophantine* (or existentially definable in terms of exponentiation) that is which possess definitions of the form:

$$(\exists u_1, \dots, u_n, v_1, \dots, v_n, w_1, \dots, w_n) [P(x_1, \dots, x_m, u_1, \dots, u_n, v_1, \dots, v_n, w_1, \dots, w_n) = 0 \\ \text{ \& } u_1 = v_1^{w_1} \text{ \& } \dots \text{ \& } u_n = v_n^{w_n}].$$

In particular, the functions $({}_k^n)$ and $n!$ were shown by her to be exponential

Diophantine. This is really what is shown in proving (1) and (2) of Theorem 4.1. The present proof of (2) is just hers; the proof of (1) given here is a simplified variant of that in [27]. (It is due independently to Julia Robinson and Matiyasevič.)

The idea of using the Chinese remainder theorem to code the effect of a bounded universal quantifier first occurred in the work of myself and Putnam [7]. In [8], we refined our methods and were able to show, beginning with the Davis normal form, that *IF* there are arbitrarily long arithmetic progressions consisting entirely of primes (still an open question), then every recursively enumerable set is exponential Diophantine. In our proof we needed to establish that $h(a, b, y) = \prod_{k=1}^y (a + bk)$ is exponential Diophantine, which we did extending Julia Robinson's methods. (The proof given here of (3) of Theorem 4.1 is a much simplified argument found much later by Julia Robinson—cf. [29].) Julia Robinson then showed first how to eliminate the hypothesis about primes in arithmetic progression, and then how to greatly simplify the proof along the lines of Lemma 5.2 of this article. Thus we obtained the theorem of [9] that every recursively enumerable set is exponential Diophantine.

Attention was now focused on the Julia Robinson hypothesis since it was plain that it would imply that Hilbert's tenth problem was unsolvable.

Many interesting propositions were found to imply the Julia Robinson hypothesis.⁷ However the hypothesis seemed implausible to many, especially because it was realized that an immediate and surprising consequence would be the existence of an absolute upper bound for the dimensions of Diophantine sets (cf. Theorem 7.6). Thus in his review [19] Kreisel said concerning the results of [9]: "... it is likely the present result is not closely connected with Hilbert's tenth problem. Also it is not altogether plausible that all (ordinary) Diophantine problems are uniformly reducible to those in a fixed number of variables of fixed degree... ."

The Julia Robinson hypothesis was finally proved by Matiyasevič [23], [24]. Specifically he showed that if we define

$$a_1 = a_2 = 1, \quad a_{n+1} = a_n + a_{n-1}$$

so that a_n is the n th Fibonacci number, then the function a_{2n} is diophantine. Then since, for $n \geq 3$, as is easily seen by induction,

$$\left(\frac{5}{4}\right)^n < a_n < 2^{n-1},$$

the set

$$D = \{\langle u, v \rangle \mid v = a_{2u} \text{ \& } u \geq 2\}$$

satisfies the Julia Robinson hypothesis. Subsequently, direct diophantine definitions of the exponential function were given by a number of investigators, several of them using the Pell equation as in this article (cf. [3], [4], [14], [18a]). The treatment in §2, 3 is based on Matiyasevič's methods, although the details are Julia Robinson's.

In particular, it was Matiyasevič who taught us how to use results like Lemmas 2.11, 2.12, and 2.22 of the present exposition. (Matiyasevič himself used analogous results for the Fibonacci numbers.)

It was soon noticed (by S. Kochen) that by a simple inductive argument the use of the Davis normal form could now be entirely avoided, as has been done in the present exposition.

Let $\#(P)$ be the number of solutions of the Diophantine equation $P = 0$. Thus $0 \leq \#(P) \leq \aleph_0$. Hilbert's tenth problem seeks an algorithm for deciding of a given P whether or not $\#(P) = 0$. But there are many related questions: Is there an algorithm for testing whether $\#(P) = \aleph_0$, or $\#(P) = 1$, or $\#(P)$ is even? I was able to show easily (beginning with the unsolvability of Hilbert's tenth problem) that all of these problems are unsolvable. In fact if

$$A = \{0, 1, 2, 3, \dots, \aleph_0\}$$

and $B \subseteq A$, $B \neq \emptyset$, $B \neq A$, then one can readily show that there is no algorithm for determining whether or not $\#(P) \in B$ (cf. [15]).

The fact that no general algorithm such as Hilbert demanded will be forthcoming adds to the interest of algorithms for dealing with special classes of Diophantine equations. Alan Baker and his coworkers [1], [2] have in recent years made considerable progress in this direction.

Notes

1. These pairing functions (but of course not their being Diophantine) were used by Cantor in his proof of the countability of the rational numbers. J. Roberts and D. Siefkes each corrected an error in the definition of these functions. They, as well as W. Emerson, M. Hausner, Y. Matiyasevič, and Julia Robinson made helpful suggestions.

2. For example, cf. [25], pp. 175–180. Matiyasevič used instead the equations $x^2 - xy - y^2 = 1$, $u^2 - muv + v^2 = 1$.

3. The recursive functions are usually defined on the nonnegative integers. This creates a minor but annoying technical problem in comparing the present definition with one in the literature (e.g., cf. [6], p. 41; also Theorem 4.2 on p. 51). Thus one can simply note that $f(x_1, \dots, x_n)$ is recursive in the present sense if and only if $f(t_1 + 1, \dots, t_n + 1) - 1$ is recursive in the usual sense. From the point of view of the intuitive "computability" of the functions involved this doesn't matter at all; one is simply in the position of using the positive integers as a "code" for the nonnegative integers — using $n + 1$ to represent n .

4. Inclusion of $S(i, u)$ in this list is redundant. That is, $S(i, u)$ can be obtained using our three operations from the remaining initial functions.

5. The method of proof is Julia Robinson's, [28], [30]. If one were permitted to use the enumeration theorem in recursive function theory ([6], p. 67. Theorem 1.4), the Universality Theorem would follow at once from Theorem 6.1.

6. Actually the result which Gödel stated (as opposed to what can be obtained at once by use of his techniques) was somewhat weaker. Indeed, the very definition of the class of recursive functions and the perception of their significance came several years later in the work of Gödel, Church, and Turing. In particular the suggestion that recursiveness was a precise equivalent of the intuitive

notion of being computable by an explicit algorithm was made independently by Church and by Turing. And of course it is this identification which is essential in regarding the technical results discussed in this account as constituting a negative solution of Hilbert's tenth problem. (For further discussion and references, cf. [6].)

7. For example, I showed ([13]) that the Julia Robinson hypothesis would follow from the non-existence of nontrivial solutions of the equation

$$9(u^2 + 7v^2)^2 - 7(x^2 + 7y^2)^2 = 2.$$

The methods used readily show that the same conclusion follows if the equation has only finitely many solutions. Čudnovskii [4] claims to have proved that 2^x is diophantine (and hence the Julia Robinson hypothesis) using this equation. Apparently there is a possibility that some of Čudnovskii's work may have been done independently of Matiyasevič — but I have not been able to obtain definite information about this.

References

1. Alan Baker, Contributions to the theory of Diophantine equations: I. On the representation of integers by binary forms, II. The Diophantine equation $y^2 = x^3 + k$, Philos. Trans. Roy. Soc. London Ser. A, 263 (1968) 173–208.
2. Alan Baker, The Diophantine equation $y^2 = ax^3 + bx^2 + cx + d$, J. London Math. Soc., 43 (1968) 1–9.
3. G. V. Čudnovskii, Diophantine predicates (Russian), Uspehi Mat. Nauk, 25 (1970) no. 4 (154), pp. 185–186.
4. ———, Certain arithmetic problems (Russian), Ordena Lenina Akad. Ukrains. SSR, Preprint IM-71–3.
5. Martin Davis, Arithmetical problems and recursively enumerable predicates, J. Symbolic Logic, 18 (1953) 33–41.
6. ———, Computability and Unsolvability, McGraw Hill, New York, 1958.
7. Martin Davis and Hilary Putnam, Reduction of Hilbert's tenth problem, J. Symbolic Logic, 23 (1958) 183–187.
8. ——— and ———, On Hilbert's tenth problem, U. S. Air Force O. S. R. Report AFOSR TR 59–124 (1959), Part III.
9. Martin Davis, Hilary Putnam, and Julia Robinson, The decision problem for exponential Diophantine equations, Ann. Math., 74 (1961) 425–436.
10. Martin Davis, Applications of recursive function theory to number theory, Proc. Symp. Pure Math., 5 (1962) 135–138.
11. ———, Extensions and corollaries of recent work on Hilbert's tenth problem, Illinois J. Math., 7 (1963) 246–250.
12. Martin Davis and Hilary Putnam, Diophantine sets over polynomial rings, Illinois J. Math., 7 (1963) 251–256.
13. Martin Davis, One equation to rule them all, Trans. New York Acad. Sci., Series II, 30 (1968) 766–773.
14. ———, An explicit Diophantine definition of the exponential function, Comm. Pure Appl. Math. 24 (1971) 137–145.
15. ———, On the number of solutions of Diophantine equations, Proc. Amer. Math. Soc., 35 (1972) 552–554.
16. Kurt Gödel, Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I, Monatsh. Math. und Physik, 38 (1931) 173–198. English translations: (1) Kurt, Gödel, On Formally Undecidable Propositions of Principia Mathematica and Related Systems, Basic

Books, 1962. (2) Martin Davis (editor), *The Undecidable*, Raven Press, 1965, pp. 5–38. (3) Jean Van Heijenoort (editor), *From Frege to Gödel*, Harvard University Press, 1967, pp. 596–616.

17. G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Fourth edition, Oxford University Press, 1960.

18. David Hilbert, *Mathematische Probleme*, Vortrag, gehalten auf dem internationalen Mathematiker-Kongress zu Paris 1900. *Nachrichten Akad. Wiss. Göttingen, Math.-Phys. Kl.* (1900) 253–297. English translation: *Bull. Amer. Math. Soc.*, 8 (1901–1902) 437–479.

18a. N. K. Kosovskii, On Diophantine representations of the solutions of Pell's equation (Russian), *Zap. Nauch. Sem. Leningrad. Otdel. Mat. Inst. Steklova*, 20 (1971) 49–59.

19. Georg Kreisel, Review of [9]. *Mathematical Reviews*, 24 (1962) Part A, p. 573 (review number A 3061).

20. Yuri Matiyasevič, The relation of systems of equations in words and their lengths to Hilbert's tenth problem (Russian). *Issledovaniya po Konstruktivnoi Matematike i Matematicheskoi Logike* II. Vol. 8, pp. 132–144.

21. ———, Two reductions of Hilbert's tenth problem (Russian), *Ibid.*, pp. 145–158.

22. ———, Arithmetic representation of exponentiation (Russian). *Ibid.*, pp. 159–165.

23. ———, Enumerable sets are Diophantine (Russian), *Dokl. Akad. Nauk SSSR*, 191 (1970) 279–282. Improved English translation: *Soviet Math. Doklady*, 11 (1970) 354–357.

23a. ———, Diophantine representation of the set of prime numbers (Russian). *Dokl. Akad. Nauk SSSR*, 196 (1971) 770–773. Improved English translation with Addendum: *Soviet Math. Doklady*, 12 (1971) 249–254.

23b. ———, Diophantine representation of recursively enumerable predicates, *Proc. Second Scandinavian Logic Symp.*, editor, J. E. Fenstad, North-Holland, Amsterdam, 1971.

23c. ———, Diophantine representation of recursively enumerable predicates, *Proc. 1970 Intern. Congress Math.*, pp. 234–238.

24. ———, Diophantine representation of enumerable predicates (Russian), *Izv. Akad. Nauk SSSR, Ser. Mat.* 35 (1971) 3–30.

24a. ———, Diophantine sets (Russian), *Uspehi Mat. Nauk*, 27(1972) 185–222.

25. Ivan Niven and Herbert Zuckerman, *An Introduction to the Theory of Numbers*, 2nd ed., Wiley, New York, 1966.

26. Hilary Putnam, An unsolvable problem in number theory, *J. Symb. Logic*, 25 (1960) 220–232.

27. Julia Robinson, Existential definability in arithmetic, *Trans. Amer. Math. Soc.*, 72 (1952) 437–449.

28. ———, Diophantine decision problems, *MMA Studies in Mathematics*, 6 (1969) [Studies in Number Theory, edited by W. J. LeVeque, pp. 76–116].

29. ———, Unsolvability of Diophantine problems, *Proc. Amer. Math. Soc.*, 22 (1969) 534–538.

30. ———, Hilbert's tenth problem, *Proc. Symp. Pure Math.*, 20 (1969) 191–194.

31. Raphael M. Robinson, Arithmetical representation of recursively enumerable sets, *J. Symb. Logic*, 21 (1956) 162–186.

32. ———, Some representations of Diophantine sets, *J. Symb. Logic*, forthcoming.

HISTORY OF THE RIEMANN MAPPING THEOREM

J. L. WALSH, University of Maryland

The Riemann mapping theorem, that *an arbitrary simply connected region of the plane can be mapped one-to-one and conformally onto a circle*, first appeared in the Inaugural dissertation of Riemann (1826–1866) in 1851. The theorem is important, for by it a result proved for the circle can often be transformed from the circle to a more general region. The proof is difficult, as involving both behavior of a function in the small (conformal mapping) and behavior in the large (one-to-one mapping). Riemann's proof was open to criticism and in the following decades numerous mathematicians sought for a proof, e.g., H. A. Schwarz (1843–1921), A. Harnack (1851–1888), H. Poincaré (1854–1912), etc., until the first rigorous proof was given in 1900 by W. F. Osgood. The proof of Osgood represented, in my opinion, the “coming of age” of mathematics in America. Until then, numerous American mathematicians had gone to Europe for their doctorates, or for other advanced study, as indeed did Osgood. But the mathematical productivity in this country in quality lagged behind that of Europe, and no American before 1900 had reached the heights that Osgood then reached.

William Fogg Osgood (1864–1943) was born in Boston in 1864, graduated from Harvard College in 1886, stayed in Cambridge for a year of graduate work, and then went to Göttingen with a Harvard fellowship for further study, especially with Felix Klein (1849–1925). According to gossip, Osgood became so enamored of a Göttingen lady that his work suffered and Klein sent him to Erlangen for his doctorate. In any case, he was accorded the degree from Erlangen in 1890 for a thesis on Abelian integrals, and one or two days later he married the girl in Göttingen, and one or two days still later they sailed for the United States of America. His

Professor Walsh received his Harvard Ph. D. under Maxime Bôcher and George David Birkhoff. He continued at Harvard as Instructor through Perkins, Professor of Mathematics and became Professor Emeritus in 1966; since then he has been at the Univ. of Maryland. He has spent leaves of absence at the Sorbonne, the Univ. of Munich, the Institute for Advanced Study, and has spent several sabbatical leaves in Paris and Jerusalem.

He is a Fellow of the American Academy of Arts and Sciences and a Member of the National Academy of Sciences. Both the SIAM Journal on Numerical Analysis and the Journal of Approximation Theory have dedicated volumes to Joseph Walsh. His main research is on zeros, extremal problems, and approximations by polynomials and orthogonal functions. He is widely known for his invention of the Walsh functions.

His publications include *Interpolation and Approximation* (Amer. Math. Soc. Coll. Series, 1935, Russian Translation — 1961), *Location of Critical Points of Analytic and Harmonic Functions* (Amer. Math. Soc. Coll. Series, 1950), *Approximation by Polynomials* (Paris, 1935), *Approximation by Bounded Analytic Functions* (Paris 1960). *The Theory of Splines and their Application* (with J. H. Ahlberg and E. N. Nilson, Academic Press, New York, 1967), *A Bibliography on Orthogonal Polynomials* (with J. S. Shohat and Einar Hille, National Research Council, Bulletin, Washington, D. C. 1940), and *A Rigorous Treatment of Maximum-Minimum Problems in the Calculus* (Heath, 1962). *Editor*.

early mathematical work was also of high quality. During the 1890's he was Lebesgue's forerunner in the study of sequences of functions of a real variable. Osgood taught at Harvard from 1890 until his retirement in 1933.

Osgood seems not to have received the recognition for his work that he deserves. For instance, C. Carathéodory and G. Julia each wrote a book on conformal mapping without mention of the name of Osgood.

We proceed now with the proof of Riemann's theorem!

By a **simply connected region** Riemann understood a region bounded by a simple closed curve, and before him special mappings by simple functions were well known. We assume the given region to be bounded, which may require an elementary preliminary transformation. Let us examine Riemann's proof (based on Dirichlet's Principle) and postpone discussion of its validity.

Mapping of a region T onto a circle is equivalent to the existence of **Green's function for T** , namely a function $G(z)$ such that

- (1) $G(z)$ is harmonic in T except at the origin 0 , assumed interior to T ;
- (2) in the neighborhood of 0 the function takes the form $G(z) \equiv G_1(z) + \log r$, where $r = |z|$ and $G_1(z)$ is harmonic throughout T ;
- (3) $G(z)$ is continuous and equal to zero at every point of the boundary C of T .

These three conditions determine $G(z)$ uniquely. Green's function for a region T is invariant under one-to-one conformal mapping of T .

If the function $w = \phi(z)$ maps T (Figure 1) onto $|w| < 1$ so that $\phi(0) = 0$, then we clearly have

$$\phi(z) = e^{G(z) + iH(z)}$$

where $H(z)$ is conjugate to $G(z)$ in T , for each of the conditions (1), (2), (3), is satisfied by $G(z)$ as thus defined. Conversely, if $G(z)$ is Green's function for T with pole in 0 , then every point of T is transformed by $w = \phi(z)$ into a point $|w| < 1$. Each locus $L_r: |\phi(z)| = r$, $0 < r < 1$ in T bounds two subregions of T , where $G(z) > \log r$ and $G(z) < \log r$ respectively; the locus L_r has no multiple points and

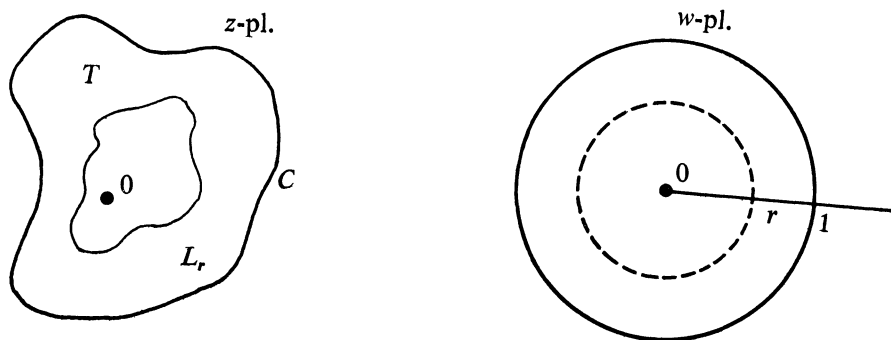


FIG. 1

separates 0 and C . On L_r we have $\partial G/\partial n \neq 0$, where n is the interior normal for the latter subregion, whence

$$\int_{L_r} \frac{\partial G}{\partial n} ds = \arg w|_{L_r} = \int_{L_r} \frac{\partial H}{\partial s} ds = \int_{L_r} \frac{\partial \log r}{\partial n} ds = 2\pi,$$

so the transformation $w = \phi(z)$ defines a one-to-one map of T onto $|w| < 1$.

If T is given, the determination of $G(z)$ requires the solution of the Dirichlet problem for T with the prescribed boundary values $\log r$ on C , a problem that Riemann treated by means of Dirichlet's principle. The physical evidence for the existence of $G(z)$ is great, for in the steady two-dimensional flow of heat, the temperature is a harmonic function provided T is a uniform body whose continuous boundary temperatures on C are prescribed.

The Dirichlet integral defined for a function $u(x, y)$ given in a region T is defined as

$$D(u) = \iint_T \left[\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 \right] dx dy \quad (\geq 0).$$

We compare this integral with the corresponding integral where $u(x, y)$ is replaced by $u(x, y) + \varepsilon \cdot v(x, y)$, where $v(x, y)$ vanishes on the boundary C of T . Thus we have, to study the function $u(x, y)$ with given boundary values minimizing $D(u)$,

$$\begin{aligned} D(u + \varepsilon v) = & \iint_T \left[\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 \right] dx dy + 2\varepsilon \iint_T \left(\frac{\partial u}{\partial x} \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \frac{\partial v}{\partial y} \right) dx dy \\ & + \varepsilon^2 \iint_T \left[\left(\frac{\partial v}{\partial x} \right)^2 + \left(\frac{\partial v}{\partial y} \right)^2 \right] dx dy. \end{aligned}$$

Considered as a function of ε , this second term on the right must be zero, namely,

$$\iint_T \frac{\partial}{\partial x} \left[\left(v \frac{\partial u}{\partial x} \right) + \frac{\partial}{\partial y} \left(v \frac{\partial u}{\partial y} \right) \right] dx dy - \iint_T v \nabla^2 u dx dy = 0$$

for all choices of the arbitrary function v . The former of these two integrals reduces to two contour integrals over C with $v (= 0$ on $C)$ as a factor of the integrand. Thus for the function u minimizing $D(u)$, $\nabla^2 u = 0$ throughout T , and u is harmonic in T . "The function solving the boundary value problem is the function minimizing $D(u)$."

This "proof," although accepted by Riemann, is obviously open to various objections:

(1) The treatment has a meaning only if C has certain properties of smoothness and differentiability.

(2) The fact that $D(u)$ has a non-negative greatest lower bound does not show the existence of a *minimum* (Weierstrass).

(3) The fact that $D(u) < \infty$ for some $u(x, y)$ satisfying the given boundary values needs to be shown (Prym 1871, Hadamard 1906).

It is convenient to assume that T is bounded; if not, we may use the transformation $w = \sqrt{(z - \alpha)/(z - \beta)}$, where α and β are two distinct boundary points of T . Then T in the z -plane corresponds to two regions T_1 and T_2 on the w -sphere, one-to-one conformal images of T , which have no common point. If two such regions do not exist, a point w_1 in T_1 can be joined to a point w_1 by a path in T_1 separating $w = 0$ and $w = \infty$, so there is a closed curve in T separating α and β , and T is not simply connected. Inversion of a point of T_2 to infinity now maps T_1 onto a bounded region.

We mention here several results that we shall need for discussion of Osgood's proof.

(1) **Axel Harnack's Theorem (1887).** If a function u_n is harmonic in a region T for all sufficiently large values of n , and if u_n increases at all points of T when n increases; if furthermore at a single point of T u_n approaches a (finite) limit when n becomes infinite; then u_n converges at all points of T to a function harmonic throughout T . (It is reported that when Harnack first told Felix Klein of this theorem, the latter refused to believe its validity.)

(2) **H. A. Schwarz.** Green's function exists for a simply connected region T bounded by a finite number of analytic arcs. (Schwarz used the alternating method, due to C. Neumann.)

(3) **Lemma.** If the bounded region T contains the closure of the region T_1 , and if O lies in T_1 , then the respective Green's functions g and g_1 with poles in O for T and T_1 satisfy the inequality $g > g_1 > 0$ in T_1 . For the difference $g - g_1$ is harmonic in T_1 , and $g_1 = 0$, $g > g_1$, on the boundary of T_1 , whence $g - g_1 > 0$, $g - g_1 \neq 0$, throughout T_1 .

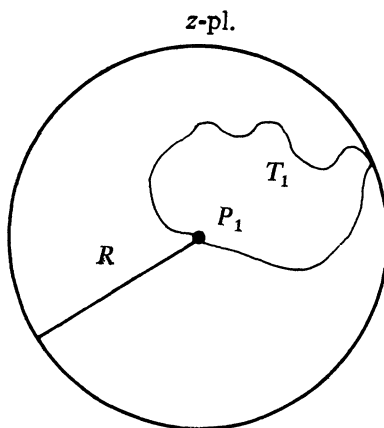


FIG. 2

(4) Given a region T_1 , it can be exhausted by a monotonic sequence of subregions, composed for instance of adjacent squares whose sides are parallel to the coordinate axes.

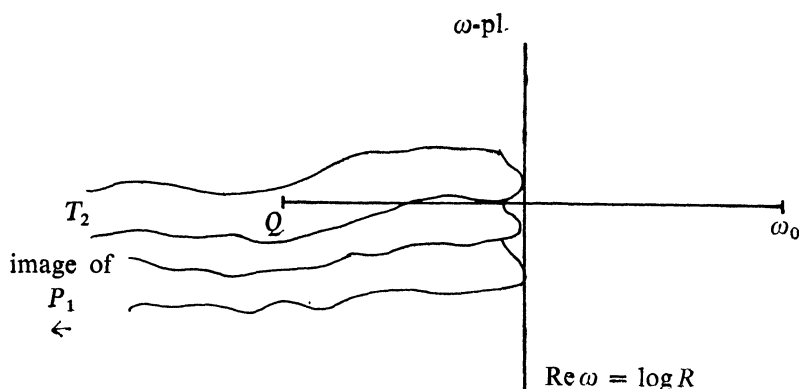


FIG. 3

Given, now, (Figure 2) a bounded simply connected region T_1 of the z -plane, we show that T_1 can be transformed into a region T of the w -plane in such a manner that a given boundary point P_1 of T_1 corresponds to a point P of a circle Γ which contains T . Let T_1 be considered to lie on the Riemann surface for $\omega = \log z$ with P_1 at $z = 0$. The image T_2 in the ω -plane of T_1 consists (Figure 3) of an infinite number of images of T_1 , each the translation of another such region by the vector

z	ω	w
P_1	∞	P
T_1	T_2	T
	ω_0	∞
	Q	O_w

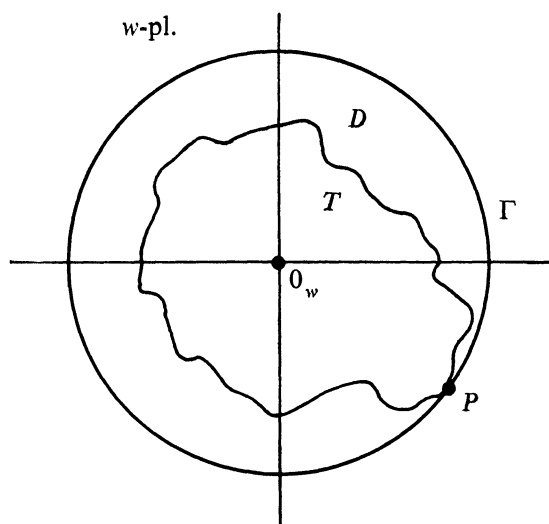


FIG. 4

$\omega = \pm 2\pi i$. In each such region the point at infinity $\omega = \infty$ corresponds to P_1 , for all boundary points of T_1 in the region $|z| < \varepsilon$ correspond to points ω with $\operatorname{Re} \omega < \log \varepsilon$. Let $|z| < R$ be the smallest circular disk with center P_1 containing T_1 ; then all points of T_2 lie in the half-plane $\operatorname{Re} \omega < \log R$. A linear transformation carrying to infinity in the w -plane a finite point ω_0 , $\operatorname{Re} \omega_0 > \log R$; carries T_2 into a region T of the w -plane (Figure 4) which lies in a circular disk D (image of $\operatorname{Re} \omega < \log R$) whose boundary passes through the image P of P_1 .

It may be noted too that an arbitrary point Q of T_2 with $\operatorname{Im} Q = \operatorname{Im} \omega_0$ can be chosen so that Q is simultaneously carried into O_w in the w -plane. The point O_w in T is then the center of D .

Let D_n be a monotonic sequence of subregions of T containing O_w , and each with a Green's function g_n with pole in O_w , with $D_{n+1} \supset D_n$, exhausting T . Let g_0 be Green's function for D with pole in O_w ; then $g_n < g_0$ in D_n . Let $g = \lim_{n \rightarrow \infty} g_n$, defined (Harnack) and harmonic throughout T except in 0. Then g is Green's function for T ; we have $0 < g_n < g_0$, $0 < g \leq g_0$. Suppose $P_k \in T$, $P_k \rightarrow P$. Since

$$\lim_{P_k \rightarrow P} g_0(P_k) = 0 \text{ for } P_k \in D,$$

we also have

$$\lim_{P_k \rightarrow P} g(P_k) = 0, \quad g(P) = 0,$$

and this shows the existence of Green's function for T and thus completes Osgood's proof of Riemann's theorem.

We have not mentioned the work of Hilbert (1862–1943), who gave a treatment of Riemann's theorem in weakened form by new methods of the Calculus of Variations, commencing about 1900. This general problem in the Calculus of Variations was presented as Problem 20 in his famous Paris lecture of 1900. He suggested in particular thesis topics on the subject for several American doctoral students in Göttingen: C. A. Noble, E. R. Hedrick, and Max Mason. However, Hilbert's method required certain smoothness properties of the boundary and of the limit function, and was thus less general than the idea of the original Dirichlet principle and less general than Osgood's proof. A new method of proof, based on function theoretic rather than potential theoretic properties, was developed by F. Riesz and L. Fejér, published in 1923 by T. Radò. Montel's theory of normal families was used, and a lemma due to Koebe. This is the standard modern proof.

Research supported in part by U. S. Air Force of Scientific Research, Grant AF 69-1690.

References

1. B. Riemann, *Grundlagen für eine allgemeine Theorie der Funktionen einer veränderlichen complexen Grösse*. Inauguraldissertation, Göttingen, 1851.

2. A. Harnack, *Logarithmisches Potential*, Leipzig, 1887, p. 167.
 3. H. A. Schwarz, *Zur Integration der partiellen Differentialgleichung $\Delta u = 0$* , *Ges. Math. Abhandlungen* vol. 11. pp. 144–171. See also Osgood, *Funktionentheorie*, Fifth Ed. Leipzig 1928, Ch. 14 § 4.
 4. Henri Poincaré, *Sur un théorème de la théorie générale des fonctions*, *Bull. Soc. Math. France*, 11 (1883) 112–125.
 5. ———, *Sur l'uniformisation des fonctions analytiques*, *Acta Mathematica*, 31 (1907) 1–63.
 6. W. F. Osgood, *On the existence of Green's function for the most general simply connected plane region*, *Trans. Amer. Math. Soc.*, 1 (1900) 310–314.
 7. ———, *Funktionentheorie*, Fifth Ed. Leipzig 1928, Ch. 14 § 5.
 8. D. Hilbert, *Über das Dirichletsche Prinzip*, *Jber. Deutsch. Math.-Verein.*, 8 (1900) 184–188.
 9. ———, *Über das Dirichletsche Prinzip*, *Math. Annalen*, 59 (1904) 161–186.
-

THE FIRST U. S. A. MATHEMATICAL OLYMPIAD

S. GREITZER, Rutgers — The State University

At its meeting on September 1, 1971, the Mathematical Association of America agreed to sponsor a U. S. A. Mathematical Olympiad in addition to the Annual High School Mathematics Examination. The purpose of the Olympiad was to attempt to discover secondary school students with superior mathematical talent — who possessed mathematical creativity and inventiveness as well as competence in computational techniques. Participation was to be limited to about 100 students selected from the Honor Roll on the High School Mathematics Examination, plus a few students of superior ability selected from those states which did not participate in the High School Mathematics Examination. The Olympiad itself was to consist of five essay-type problems requiring mathematical power on the part of the participants. The committee responsible for conducting the Olympiad consisted of Samuel L. Greitzer, Rutgers University, Alfred Kalfus, Babylon High School, Murray S. Klamkin, Ford Motor Company, and Nura D. Turner, SUNY at Albany.

Invitations were sent to 106 students on April 14, 1972, and 100 students took the Olympiad on May 9, 1972. The committee which prepared the Olympiad consisted of Murray Klamkin, D. J. Newman, Yeshiva University and Abraham Schwartz, CUNY. The Olympiad is reproduced below. (Solutions have been provided at the end of this article.)

THE FIRST U. S. A. MATHEMATICAL OLYMPIAD

MAY 9, 1972

1. The symbols (a, b, \dots, g) and $[a, b, \dots, g]$ denote the greatest common divisor and the least common multiple, respectively, of the positive integers a, b, \dots, g . For

example, $(3, 6, 18) = 3$ and $[6, 15] = 30$. Prove that

$$\frac{[a, b, c]^2}{[a, b][b, c][c, a]} = \frac{(a, b, c)^2}{(a, b)(b, c)(c, a)}.$$

2. A given tetrahedron $ABCD$ is isosceles, that is, $AB = CD$, $AC = BD$, $AD = BC$. Show that the faces of the tetrahedron are acute-angled triangles.

3. A random number selector can only select one of the nine integers 1, 2, ..., 9, and it makes these selections with equal probability. Determine the probability that after n selections ($n > 1$), the product of the n numbers selected will be divisible by 10.

4. Let R denote a nonnegative rational number. Determine a fixed set of integers a, b, c, d, e, f such that, for every choice of R ,

$$\left| \frac{aR^2 + bR + c}{dR^2 + eR + f} - \sqrt[3]{2} \right| < \left| R - \sqrt[3]{2} \right|.$$

5. A given convex pentagon $ABCDE$ has the property that the area of each of the five triangles ABC , BCD , CDE , DEA and EAB is unity. Show that every non-congruent pentagon with the above property has the same area, and that, furthermore, there are an infinite number of such noncongruent pentagons.

All 100 participants mailed in their papers by May 15, and they were graded by a committee consisting of Dr. John Bender, Dr. Richard Bumby, Dr. L. M. Kelly, Dr. Sol Leader, Dr. B. Muckenhoupt and Dr. H. Zimmerberg. We are indebted to the mathematics department of Rutgers University for their help in grading. The final results were sent out to the participating schools on June 1, 1972. These results are tabulated below:

H.S.Exam Olymp.	85— 89.75	90— 94.75	95— 99.75	100— 104.75	105— 109.75	110— 114.75	115— 119.75	120— 124.75	125 129.7
90-99						1			
80-89				1			1		
70-79					1	3	1		
60-69	1				1				
50-59	6	2	6	1	1				1
40-49	7	1		1	1	1			
30-39	2	5	4	1	1	1	1		
20-29	5	4	2	1	1	1			
10-19	5	1	1	3				1	
0-9	6	3	2	1	1			2	

The eight finalists selected to receive awards had scores indicated in the rectangle at the upper part of the table.

While it is not safe to draw conclusions on the basis of so small a sample, it would appear that (a) a high score on the High School Mathematics Examination did not necessarily lead to a high score on the Olympiad. In fact, the correlation is only 0.24. We can conclude that the Olympiad is measuring something other than does the High School Mathematics Examination. However, (b) it must be noted that all eight finalists did score above 100 on the High School Mathematics Examination. We may conclude that students rated superior on the Olympiad are also superior on the High School Mathematics Examination, but that the converse is not necessarily true.

Interest in the Olympiad was high. Members of the Olympiad Committee received numerous calls and letters asking for further information, offers to help, and pleas to be allowed to take part. On May 8, we received calls from two schools that had not received the contest materials. In both cases, the difficulty originated with the school. It was too late, however, to mail anything to the schools. Therefore, Mr. John Clark drove to Rutgers from Garden City, New York, about 75 miles each way, and Mr. Kazlouskas drove to Rutgers from Binghamton, New York, about 200 miles each way, and picked up the Olympiad materials personally. Both supervisors are to be commended for their interest in their students. Incidentally, the student from Garden City was tied for second place. Plans had originally called for special recognition for the top five contestants. Because there were ties for second, third and fourth places, the number of finalists was increased to eight. The list of finalists is given below:

1	James Saxe	Albany High School	Albany, N. Y.
2	{ Thomas Hemphill David Vanderbilt	James Monroe High School Garden City High School	Sepulveda, Calif. Garden City, N. Y.
3	{ Paul Harrington Arthur Rubin	Paul V. Moore H. S. West Lafayette H. S.	Central Square, N. Y. West Lafayette, Ind.
4	{ David Anick Steven Rahe	Ranney School Central High School	New Shrewsbury, N. J. Sioux City, Iowa
5	James Shearer	Livermore High School	Livermore, Calif.

The problem of providing suitable recognition for these very talented students was solved for the Committee through the generosity of International Business Machines, Inc., which agreed to supply funds for a ceremony. Dr. Nura Turner was in charge of this ceremony, which consisted of two days of activities in Washington, D. C. All eight contestants were present, some with parents (traveling at their own expense and paying for their own accommodations).

On Tuesday, September 12, there was a reception at the National Academy of Sciences, where the finalists were given awards presented by the Olympiad Committee

and the NCTM. Dr. Emanuel Piore spoke, congratulating the finalists and discussing the roles of mathematicians in science and mathematics. This was followed by a dinner at the Department of State.

On Wednesday, September 13, the group toured the White House, and were greeted by Dr. Edward E. David, Science Advisor to the President. They then visited the National Bureau of Standards, had lunch there, after a greeting on behalf of Dr. Burton Colvin. In the afternoon, they listened to lectures on mathematical applications by Dr. M. Newman, Dr. Wesley Nicholson, Dr. R. A. Kirsch and Dr. A. J. Goldman.

The Committee acknowledges with thanks the help provided by these organizations, and hopes they will be equally helpful to future Olympiad finalists.

The Olympiad Committee:

S. Greitzer (chairman)
A. Kalfus
M. Klamkin
N. Turner

Solutions to Problems

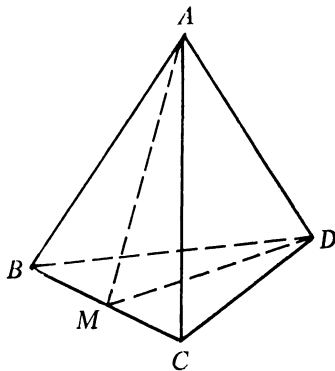
1. Let $a = \prod p_i^{a_i}$, $b = \prod p_i^{b_i}$, $c = \prod p_i^{c_i}$ where p_i denote the prime factors of a, b, c (some of the exponents may be zero).

Since $[a, b] = \prod p_i^{\max(a_i, b_i)}$ $(a, b) = \prod p_i^{\min(a_i, b_i)}$ etc., we have to show that

$$\begin{aligned} & 2 \max(a_i, b_i, c_i) - \max(a_i, b_i) - \max(b_i, c_i) - \max(c_i, a_i) \\ &= 2 \min(a_i, b_i, c_i) - \min(a_i, b_i) - \min(b_i, c_i) - \min(c_i, a_i). \end{aligned}$$

Without loss of generality, let $a_i \geq b_i \geq c_i$ for any particular index i . Then $2a_i - a_i - b_i - a_i = 2c_i - b_i - c_i - c_i$. (One contestant gave analogous results for four and five numbers.)

2. It follows immediately that the sum of the face angles at any vertex is 180° . Since the sum of any two angles of a trihedral angle is greater than the third, each



angle must be acute. Again (see diagram) assume that $\triangle BDC$ is nonacute. Let M be the midpoint of BC . Since $\triangle ABC \cong \triangle DCB$, $AM = DM$. Now consider a circle with M as center and BC as a diameter. Since $\angle BDC$ is nonacute, C must lie within or on the circle. Thus, $2DM \leq BC$. By the triangle inequality on $\triangle AMD$, $AM + MD > AD$. But since $AM = MD$ and $AD = BC$, we obtain a contradiction. Therefore, $\angle BDC$ is acute.

3. In order for the product to be divisible by 10, there must be at least one 5 and at least one even number among the n numbers that have come up. Let A denote the event of obtaining at least one 5 and let B denote the event of obtaining at least one even number in n spins. If $A + B$, AB , A' , $P(E)$ denote the union of a , b ; the intersection of A, B ; the complement of A ; the probability of the event E , respectively, we have $P(AB) = 1 - P(A') - P(B') + P(A'B')$. Hence

$$P(AB) = 1 - \left(\frac{8}{9}\right)^n - \left(\frac{5}{9}\right)^n + \left(\frac{4}{9}\right)^n.$$

4. Given the inequality, we wish to determine fixed a, b, c, d, e, f to satisfy it, for all nonnegative rational R . As $R \rightarrow \sqrt[3]{2}$ through a sequence of rational numbers, the right hand side of this inequality approaches zero. Consequently, the left hand side must vanish if we set $R = \sqrt[3]{2}$. Therefore,

$$a \cdot 2^{2/3} + b \cdot 2^{1/3} + c = 2d + e \cdot 2^{2/3} + f \cdot 2^{1/3}.$$

It is then *necessary* that $a = e$, $b = f$, $c = 2d$. On substituting back into the inequality and factoring out the common factor $R - \sqrt[3]{2}$ from both sides, we obtain,

$$\left| \frac{aR + b - d \cdot 2^{1/3}(R + 2^{1/3})}{dR^2 + aR + b} \right| < 1.$$

For the last inequality to be satisfied, it *suffices* to let a, b, d be positive integers and make the numerator nonnegative, i.e., by letting $a > d \cdot 2^{1/3}$, $b > d \cdot 4^{1/3}$. Two simple choices are $d = 1$, $a = b = 2$, and $d = 3$, $a = 4$, $b = 5$ leading respectively to

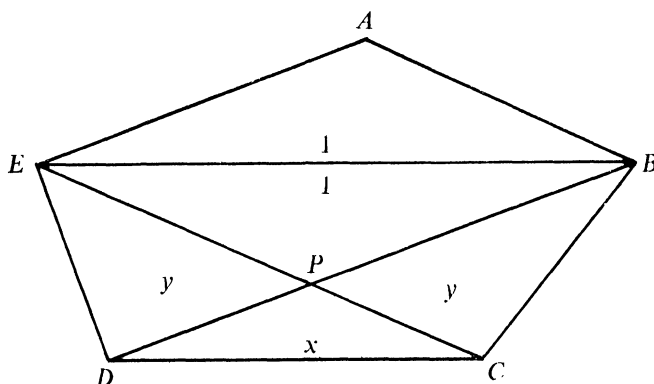
$$\frac{2R^2 + 2R + 2}{R^2 + 2R + 2} \quad \text{and} \quad \frac{4R^2 + 5R + 6}{3R^2 + 4R + 5}.$$

It is to be noted that the second approximation is a better one than is the first.

5. Since $\triangle EDC = \triangle BDC = 1$, both triangles have equal altitudes to side CD . Thus $DC \parallel EB$. Similarly the other diagonals are parallel to their respective opposite sides. Whence, $ABPE$ is a parallelogram and $\triangle PEB = 1$. Letting $\triangle EDP = y = \triangle PBC$ and $\triangle PDC = x$, we then have $x + y = 1$. Also:

$$\frac{\triangle EPD}{\triangle EPB} = \frac{DP}{PB} = \frac{\triangle DPC}{\triangle CPB} \quad \text{or} \quad \frac{y}{1} = \frac{x}{y}.$$

Thus $y^2 + y - 1 = 0$ and $y = (\sqrt{5} - 1)/2$. Area $(ABCDE) = 2 - y - x - y = (5 + \sqrt{5})/2$.



To prove that there is an infinite number of such noncongruent pentagons, construct an arbitrary triangle PDC whose area is x . Extend CP to E and DP to B , so that $\triangle EDC = \triangle BDC = 1$. Now draw $EA \parallel BD$ and $AB \parallel EC$. It follows from the previous analysis that the pentagon has the desired property.

Another proof of the last part can be gotten easily by parallel projection of a regular pentagon with the desired property, i.e., let $x' = mx$, $y' = y/m$ (m arbitrary). Under this transformation, areas are preserved. (One contestant showed that these projections constitute *all* the solutions to the problem.)

CORRECTION TO "WHAT IS A RECIPROCITY LAW?"

This MONTHLY, 79:571-586 (June-July, 1972)

B. F. WYMAN, Ohio State University

I wish to thank Karl A. Beres and Lawrence J. Dickson for pointing out that the theorem of Section 6 (p. 583) is incorrect. Professor Dickson has derived correct formulas for the v_m and μ_j (defined on p. 582):

$$(1) \quad v_m = \sum_{j=1}^n (j, m) \mu_j, \quad n = \deg(f).$$

$$(2) \quad \mu_j = \sum_{r=1}^{[n/j]} \sum_{m|jr} (\mu(r) \mu(m) / \phi(jr)) v_{jr/m},$$

where ϕ and μ are the Euler and Möbius functions.

Since the theorem referred to was not used in the preparation of the computer program mentioned in Section 7, the numerical results reported there are unaffected.

MATHEMATICAL NOTES

EDITED BY ROBERT GILMER

The present backlog for this Department is substantial. Until further notice, new manuscripts cannot be accepted. This moratorium will probably continue until June 1, 1973; authors are requested to hold their manuscripts pending a further announcement.

ON ELEMENTARY PROOFS OF PEANO'S EXISTENCE THEOREMS

JOHANN WALTER, Technische Hochschule, 51 Aachen, FRG.

1. Historical and didactical remarks. The theory of ordinary differential equations is an important building stone in the curriculum of every student of applied mathematics. The fundamental problems arising in this theory have attracted and deserve persevering attention, particularly from the didactical point of view. (For a recent discussion of these questions cf. [14].) The nonexistence even for very familiar differential equations—like certain Sturm-Liouville or Riccati equations—of “closed formulas” representing the solutions of these equations in terms of “elementary functions” is one of the most puzzling phenomena we are confronted with in this part of analysis. It is by this phenomenon that we are forced to look for existence theorems of a general kind. The most important of these existence theorems as regards the weakness of its assumptions as well as the conceptual simplicity of the idea underlying its proof is the existence theorem of Peano. The didactical requirement thus arises to have a proof as simple and as elementary as possible of this theorem or of a significant special case of it.

In this connection Kennedy [6] has proposed as a “Research Problem” the question “Is there an elementary proof of Peano’s existence theorem for first order differential equations?” Before entering into a more detailed discussion of this Research Problem we have to clear up two terminological questions.

1.) Kennedy calls attention to two papers [9], [10] (reprinted in [11], p. 74 and 119) of Peano. In these papers Peano formulates two existence theorems differing from each other. Consider the ordinary differential equation

$$x' = f(t, x),$$

where x and f belong to Euclidean d -space R^d and t is real. Let K be a positive number, $J = [0, 1]$, f continuous in $J \times R^d$ and $|f(t, x)| \leq K$ for $(t, x) \in J \times R^d$.

THEOREM 1 ([9], 1886). *Let $d = 1$. The initial value problem*

$$(1) \quad x'(t) = f(t, x(t)) \text{ for } t \in J, \quad x(0) = 0,$$

has solutions x_{\min} , x_{\max} such that $x_{\min}(t) \leq x(t) \leq x_{\max}(t)$ for $t \in J$, holds for every solution $x(\cdot)$ of (1).

THEOREM 2 ([10], 1890). *Let $d \geq 1$. The initial value problem (1) has at least one solution.*

REMARK. In the case $d = 1$, Theorem 2 is a trivial consequence of Theorem 1. This means that every proof of Theorem 1 is also a proof of Theorem 2. On the other hand, the proofs of Theorem 2 are more simple than those of Theorem 1.

All authors who mention the name of Peano at all agree to call Theorem 2 *the* Theorem of Peano. Kamke [5] and Hille [4] do not even seem to know [9]. Now it is not completely clear to which of the two theorems Kennedy's question does refer. Although he cites Theorem 2 at the beginning of his paper, it is obviously Theorem 1 he has in view later on. (Similarly W. Walter in his reply [15] "There is an elementary proof of Peano's existence theorem" quotes Theorem 2 (for $d = 1$) but proves the stronger Theorem 1.)

2.) It can be inferred from the work of Kennedy [6] and W. Walter ([15] and *Autorreferat in Zentralblatt für Mathematik* 207 (1971), 84) that these authors call a proof elementary if it avoids equicontinuous families and an appeal to the lemma of Arzelà-Ascoli and if the special properties of R^1 (as compared with R^d) are used instead. We shall not enter into the notion of constructiveness emphasized in this connection by several authors ([3], [8], [15] but not [9]).

Do there exist proofs of Peano's theorems which are elementary in the sense just described?

Ad Theorem 1: Kennedy himself specifies three proofs of this kind: Peano's original proof [9], Perron's proof [12] and Osgood's proof [8]. But he takes the first two proofs to be defective, the third being rather complicated. W. Walter in his reply [15] gives another elementary proof and refers to the proof of Grunsky [3]. Moreover, he stresses that Perron's proof is correct (In the opinion of the present author, Peano's proof is also essentially correct!)

Ad Theorem 2: In the proof of Theorem 2 for $d > 1$ the lemma of Arzelà-Ascoli seems unavoidable. Every time, however, the special case $d = 1$ of Theorem 2 is treated separately (e.g., in [1], [2], [7], [13]) its proof is based on the lemma of Arzelà-Ascoli too.

Résumé: Although a great number of elementary proofs of Theorem 1 exist, an elementary proof of Theorem 2 in the case $d = 1$ is still lacking in the literature. A proof of this kind is of an at least didactical interest, because, as we have seen, the case $d = 1$ of Theorem 2 is treated separately in several textbooks and because, on the assumption of Theorem 2, the proof of the existence of x_{\min} , x_{\max} can be given in an especially suggestive way (cf., [1], [4]): It is possible to represent x_{\min} resp. x_{\max} as the infimum resp. supremum of the set of all solutions of (1) provided this set is not empty. But this is exactly the statement of Theorem 2.

In the following we shall give an elementary proof of Theorem 2 for $d = 1$. In

this proof (just as in [1], [5], [7], [8], [12], [15]) use is made of the integral equation

$$x(t) = \int_0^t f(s, x(s)) ds$$

associated with (1). An elementary proof without this integral equation (as in [2], [3], [4], [9], [10], [13]) would perhaps also be interesting.

2. An elementary proof of Peano's second theorem for $d = 1$. Let

$$E = \{t_0, t_1, \dots, t_N\}, \quad 0 = t_0 < t_1 < \dots < t_N = 1,$$

be a partition of J with the norm

$$|E| = \sup_{k=1, \dots, N} (t_k - t_{k-1}).$$

We define $x_0 = 0$ and $x_k = x_{k-1} + (t_k - t_{k-1}) f(t_{k-1}, x_{k-1})$ for $k = 1, \dots, N$. An approximate solution $p_E(\cdot)$ is then constructed by joining the points (t_k, x_k) by a polygonal line; p_E is called the Euler polygon associated with the partition E . By construction we have

$$(2) \quad \begin{aligned} p_E(0) &= 0, \quad |p_E(t)| \leq K \text{ for } t \in J, \\ |p_E(t_1) - p_E(t_2)| &\leq K \cdot |t_1 - t_2| \text{ for } t_1, t_2 \in J. \end{aligned}$$

The following three lemmas are also used (at least implicitly) by W. Walter [15].

LEMMA 1. *Let E_n , $n = 1, 2, \dots$, be a sequence of partitions of J such that $|E_n| \rightarrow 0$, $n \rightarrow \infty$. Then a function $N(\cdot)$ defined for positive real numbers exists such that (except at a finite number of points where p'_{E_n} does not exist)*

$$(3) \quad n > N(\varepsilon) \Rightarrow |p'_{E_n}(t) - f(t, p_{E_n}(t))| < \varepsilon.$$

LEMMA 2. *Let p and p_n , $n = 1, 2, \dots$, be continuous functions defined in J such that $p(t) = \lim_{n \rightarrow \infty} p_n(t)$ uniformly in J . Then if*

$$(4) \quad \lim_{n \rightarrow \infty} \left| p_n(t) - \int_0^t f(s, p_n(s)) ds \right| = 0 \text{ for } t \in J,$$

p is a solution of (1).

Now the usual proof of Theorem 2 ($d = 1$) proceeds as follows: Because of (2) the sequence p_{E_n} is a family of uniformly bounded equicontinuous functions. On account of the lemma of Arzelà-Ascoli there exists a subsequence converging uniformly in J . For this subsequence (3) holds a fortiori. By integrating (3) we get (4). This completes the usual proof.

The lemma of Arzelà-Ascoli has the following corollary which can also easily be proved directly.

LEMMA 3. Let x and x_n , $n = 1, 2, \dots$, be functions defined in J such that

$$(5) \quad x(t) = \lim_{n \rightarrow \infty} x_n(t) \quad \text{for every } t \in J,$$

$$(6) \quad |x_n(t_1) - x_n(t_2)| \leq K|t_1 - t_2| \quad \text{for } t_1, t_2 \in J, \quad n = 1, 2, \dots.$$

Then we have $|x(t_1) - x(t_2)| \leq K|t_1 - t_2|$ for $t_1, t_2 \in J$, x_n converges uniformly in J to x .

Using Lemma 3 the application of the lemma of Arzelà-Ascoli in the usual proof sketched above obviously can be avoided if a sequence of functions satisfying (4), (5) and (6) is available. A sequence of this kind can easily be obtained exploiting a special property of R^1 , viz., the order relation. Let E_n , $n = 1, 2, \dots$, be a sequence of partitions of J such that $|E_n| \rightarrow 0$, $n \rightarrow \infty$, and for $t \in J$ define

$$(7) \quad p_{n,k}(t) = \sup_{n \leq j \leq k} p_{E_j}(t), \quad n \leq k,$$

$$(8) \quad p_n(t) = \sup_{n \leq k} p_{n,k}(t),$$

$$(9) \quad p(t) = \inf p_n(t).$$

In the following it will be shown that the sequence p_n , $n = 1, 2, \dots$, has the properties (4), (5) and (6) required above. Firstly all functions just defined are bounded in absolute value by K . Because of (7), $p_{n,k}(t)$ is increasing in k and because of (8), $p_n(t)$ is decreasing in n .

Therefore we have

$$(10) \quad p_n(t) = \lim_{k \rightarrow \infty} p_{n,k}(t),$$

$$(11) \quad p(t) = \lim_{n \rightarrow \infty} p_n(t).$$

From (11) we infer that p_n satisfies (5). $p_{n,k}$ is a polygonal line (with only a finite number of jumps in the first derivative) and by (2) and (7) also lipschitzian with Lipschitz constant K . Thus for fixed n the polygons $p_{n,k}$ satisfy the assumptions (5), (6) of Lemma 3. Therefore the convergence of (10) is uniform and p_n is also lipschitzian with Lipschitz constant K . This means that the sequence p_n also satisfies (6). Moreover, the relation (3) remains true (uniformly in k) if p_{E_n} is replaced by $p_{n,k}$. Integrating this relation we get

$$(12) \quad n > N(\varepsilon) \Rightarrow \left| p_{n,k}(t) - \int_0^t f(s, p_{n,k}(s)) ds \right| < \varepsilon,$$

(also uniformly in k). Passing to the limit with respect to k in (12) we arrive at (4), q.e.d.

References

1. E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York-Toronto-London, 1955.
2. L. E. El'sgol'ts, *Differential Equations*, Hindustan Publishing, Delhi, 1961.
3. H. Grunsky, Ein konstruktiver Beweis für die Lösbarkeit der Differentialgleichung $y' = f(x, y)$ bei stetigem $f(x, y)$, *Jber. Deutsch. Math. - Verein. Abt. 1*, 63 (1960) 78-84.
4. E. Hille, *Lectures on Ordinary Differential Equations*, Addison-Wesley, Reading, Mass., 1969.
5. E. Kamke, *Differentialgleichungen I*, 5. Auflage, Akademische Verlagsgesellschaft Geest & Portig K.-G., Leipzig, 1964.
6. H. C. Kennedy, Is there an elementary proof of Peano's existence theorem for first order differential equations? *This MONTHLY*, 76 (1969) 1043-1045.
7. F. J. Murray and K. S. Miller, *Existence Theorems for Ordinary Differential Equations*, New York University Press, New York, 1954.
8. W. F. Osgood, Beweis der Existenz einer Lösung der Differentialgleichung $y' = f(x, y)$ ohne Hinzunahme der Cauchy-Lipschitzschen Bedingung, *Monatsh. Math.*, 9 (1898) 331-345.
9. G. Peano, Sull'integrabilità delle equazioni differenziali del primo ordine, *Atti Accad. Sci. Torino*, 21 (1886) 677-685.
10. ———, Démonstration de l'intégrabilité des équations différentielles ordinaires, *Math. Ann.*, 37 (1890) 182-228.
11. ———, *Opere Scelte*, vol. 1, edited by Ugo Cassina, Rome, 1957-1959.
12. O. Perron, Ein neuer Existenzbeweis für die Integrale der Differentialgleichungen $y' = f(x, y)$, *Math. Ann.*, 76 (1915) 471-484.
13. J. G. Petrovski, *Ordinary Differential Equations*, Prentice Hall, Englewood Cliffs, N. J., 1966.
14. A. Strauss and J. A. Yorke, On the fundamental theory of differential equations, *SIAM Review*, 11 (1969) 236-246.
15. W. Walter, There is an elementary proof of Peano's existence theorem, *this MONTHLY*, 78 (1971) 170-173.

A REMARK CONCERNING ABSOLUTELY CONTINUOUS FUNCTIONS

F. S. VAN VLECK, University of Kansas and University of Colorado

In a recent note Goffman [1] gave a short, clear proof of the well-known theorem:

THEOREM A. *A function f whose derivative f' exists everywhere and is summable is absolutely continuous.*

The purpose of this note is to point out that Goffman essentially proved somewhat more — by trivially modifying his argument one obtains a not so well-known characterization of absolutely continuous functions.

It is well known that the derivative of an absolutely continuous function exists almost everywhere. There is also a standard counterexample [cf. 2, p. 168] that everywhere existence of f' cannot be replaced by almost everywhere existence in The-

orem A. In [3, p. 183] and [4, Exercise 18.41 (d)] it is shown that f' existing everywhere can be relaxed to f' existing except on a countable set. The following theorem shows what additional condition must be assumed in order to replace everywhere existence of f' by almost everywhere existence.

THEOREM B. *Let I be a closed interval and $f: I \rightarrow \mathbb{R}$. Necessary and sufficient conditions for f to be absolutely continuous are:*

- (i) f is continuous on I .
- (ii) f' exists almost everywhere on I and is summable.
- (iii) For $E \subset I$, $\mu(E) = 0$ implies $\mu(f(E)) = 0$.

Condition (iii) is sometimes expressed by saying f is an N -function or that f has Property N . According to [5, p. 224] this concept is due to N. N. Lusin. By [1, Lemma 1], f' existing everywhere implies (iii) and so Theorem B yields Theorem A as an immediate corollary.

The necessity part of Theorem B is well known, although condition (iii) seems to be neglected in most, but not all, analysis texts. Hewitt and Stromberg [4] give a discussion of (iii) and, in particular, give another characterization of absolutely continuous functions, due to Banach and Zarecki, which involves Property N . Saks [5, pp. 224–228] also discusses this condition and explicitly states Theorem B. Varberg [6, Theorem 3] proves an n -dimensional version of the sufficiency part of Theorem B along the lines of Saks' proof.

We now show that Goffman essentially proved the sufficiency part of Theorem B. We indicate what changes must be made in his argument. First, in [1, Corollary] replace the hypothesis " f is everywhere differentiable" by " f has property (iii)". The conclusion is the same and the proof is essentially the same. Next, replace the hypotheses of [1, Lemma 2] by " f satisfies (i)–(iii)". The conclusion and proof remain as in [1]. The rest of the proof of the sufficiency in Theorem B is exactly as in [1, Proof of Theorem].

D. Varberg has kindly pointed out that these theorems also appear in his article "On absolutely continuous functions", this MONTHLY, 72 (1965) 831–841.

References

1. C. Goffman, On functions with summable derivatives, this MONTHLY, 78 (1971) 874–875.
2. W. Rudin, Real and Complex Analysis, McGraw-Hill, New York, 1966.
3. H. Kestelman, Modern Theories of Integration, Dover, New York, 1960.
4. E. Hewitt and K. Stromberg, Real and Abstract Analysis, Springer-Verlag, New York, 1965.
5. S. Saks, Theory of the Integral, Dover, New York, 1964.
6. D. Varberg, On differentiable transformations in R^n , this MONTHLY, 73 (1966) 111–114.

ON NON-ASSOCIATIVE ALGEBRAS DERIVED FROM GRAPHS

W. E. JENNER, University of North Carolina, Chapel Hill

A construction is given for associating an algebra with any finite graph in such a way that the algebra is simple if and only if the graph is connected. The algebras obtained are in general nonassociative and the simple algebras appear to be new.

A **graph**, in the sense it will be used here, is a set of objects called **vertices**, together with a collection of two-element subsets called **edges**. Two vertices belonging to a given edge are said to be **adjacent**. A graph is said to be **connected** if for any two of its vertices p and q there is a sequence of vertices, beginning with p and terminating with q , such that any two successive vertices in the sequence are adjacent. A graph is **finite** if it has a finite set of vertices.

Now let Γ be a finite graph and K be any field. Let $\mathfrak{A}(\Gamma)$ be the set of all functions of $\Gamma \times \Gamma$ into K . This is made into an algebra in the following manner. First, if $f, g \in \mathfrak{A}(\Gamma)$ and $\lambda, \mu \in K$ then $\lambda f + \mu g$ is the mapping $(p, q) \mapsto \lambda f(p, q) + \mu g(p, q)$. This makes $\mathfrak{A}(\Gamma)$ into a vector space. The elements $\{e_{pq}\}$ constitute a basis, where e_{pq} is the mapping taking the pair (p, q) into $1 \in K$ and all other pairs into $0 \in K$. Multiplication in $\mathfrak{A}(\Gamma)$ is defined in terms of the basis elements as follows: $e_{pq} \cdot e_{rs} = 0$ if $q \neq r$; $e_{pq} \cdot e_{qs} = e_{ps}$ if $p = q$, or $q = s$, or p and q are adjacent, or q and s are adjacent; otherwise the product is zero.

If Γ has n elements and any two vertices are adjacent, then $\mathfrak{A}(\Gamma)$ is isomorphic to the associative algebra $[K]_n$. In general, however, $\mathfrak{A}(\Gamma)$ will not be associative. Suppose, for instance, that p and q are adjacent, q and r are adjacent, but p and r are not adjacent. Then $(e_{rp} \cdot e_{pq}) \cdot e_{qr} = e_{rq} \cdot e_{qr} = e_{rr}$, whereas $e_{rp} \cdot (e_{pq} \cdot e_{qr}) = e_{rp} \cdot e_{pr} = 0$.

THEOREM. *If Γ is a finite connected graph, then $\mathfrak{A}(\Gamma)$ is a simple algebra.*

Proof. Suppose \mathfrak{a} is a nonzero ideal of $\mathfrak{A}(\Gamma)$ and that $f = \sum \lambda_{pq} e_{pq}$ ($\lambda_{pq} \in K$) is a nonzero element of \mathfrak{a} . Suppose $\lambda_{rs} \neq 0$. Then $(e_{rr} \cdot f) \cdot e_{ss} = \lambda_{rs} e_{rs} \in \mathfrak{a}$ and so $e_{rs} \in \mathfrak{a}$. Take any $p, q \in \Gamma$. Since Γ is connected, there exists a sequence p, a, b, \dots, k, r of vertices such that any two successive terms are adjacent. Then $e_{pa} \cdot (e_{ab} \cdot (\dots (e_{kr} \cdot e_{rs}) \dots)) = e_{ps} \in \mathfrak{a}$. Similarly, operating on the right, it follows that $e_{pq} \in \mathfrak{a}$ and so $\mathfrak{a} = \mathfrak{A}(\Gamma)$. Thus $\mathfrak{A}(\Gamma)$ is simple.

The structure of the algebra $\mathfrak{A}(\Gamma)$ can explicitly be determined for any finite graph Γ , not necessarily connected. Indeed let $\Gamma = \Gamma_1 \cup \dots \cup \Gamma_t$ be the decomposition of Γ into its connected components. Then $\mathfrak{A}(\Gamma) = \mathfrak{A}(\Gamma_1) + \dots + \mathfrak{A}(\Gamma_t) + \mathfrak{N}$, vector space direct sum, where \mathfrak{N} is the subspace of $\mathfrak{A}(\Gamma)$ spanned by those e_{pq} where p and q belong to different components Γ_i . It is verified immediately that \mathfrak{N} is a zero algebra and an ideal of $\mathfrak{A}(\Gamma)$. Now $\mathfrak{A}(\Gamma) - \mathfrak{N}$ is the direct sum of the $\mathfrak{A}(\Gamma_i)$ and so \mathfrak{N} is the radical of $\mathfrak{A}(\Gamma)$ in any reasonable sense. The sum of the

$\mathfrak{U}(\Gamma_i)$, which is direct, is a subalgebra \mathfrak{U}_0 of $\mathfrak{U}(\Gamma)$, so that there is a Wedderburn Principal Theorem $\mathfrak{U}(\Gamma) = \mathfrak{U}_0 + \mathfrak{N}$.

The author is indebted to Ladnor Geissinger, conversations with whom suggested that there might be algebras of this sort; also to Douglas Kelly for helpful advice. The work reported here was done at the Seminar on Combinatorial Theory held at Bowdoin College in the summer of 1971.

A FINITE DIFFERENCE PROOF THAT $E = mc^2$

DONALD GREENSPAN, University of Wisconsin

Abstract. It is shown that the classical formula $E = mc^2$ follows directly from forward difference definitions of both velocity and energy, thus avoiding the necessity of the concept of a derivative.

In order to reach the reader of minimal background, Taylor and Wheeler [2] developed the theory of special relativity using differences, whenever possible, rather than derivatives. In this note we shall show that the classical formula $E = mc^2$ can, in fact, be established entirely without the concept of a derivative. Such a result is not only of interest in itself, but it also affirms the intrinsic role of finite differences in the development of physical models, a result already substantiated by the application of high speed computers in solving nonlinear problems of applied science [1].

First, let us summarize in a convenient way the basic concepts which are necessary for the discussion. Consistently, we measure not only length, but also time, in the same unit, meters, as follows. A meter of time, denoted by 1 meter/ c , is the time it takes for light to travel one meter. Thus,

$$(1) \quad 1 \text{ meter}/c = (3.335640)10^{-9} \text{ sec.}$$

It is assumed that at every point in Euclidean three-space there is a clock which is synchronized with the clock at the origin. When one observes an event and records not only its position but also the time on the clock at that position, one says that an observation has been made in space-time. The coordinates of an event are of the form (x, y, z, t) . With regard to the observation of events in space-time, it will be assumed that the coordinate system is inertial and that all laws of physics are the same in every inertial reference frame.

Though time will always be measured in meters, it is sometimes convenient to measure speed conventionally as v meters per second, or in light-time as β meters per meter. Thus, if t_1 and t_2 are any two time readings such that

$$t_2 - t_1 = 1 \text{ meter}/c,$$

and if a particle in motion along an X -axis is at x_1 at time t_1 and at x_2 at time t_2 , then we define β and v at t_1 by the forward differences

$$(2) \quad \beta = \frac{x_2 - x_1}{t_2 - t_1},$$

$$(3) \quad v = \frac{x_2 - x_1}{(t_2 - t_1)(3.335640)10^{-9}}.$$

The units of β are then meters per meter, while the units of v are meters per second. From (2) and (3) one has

$$(4) \quad \beta = v/c.$$

Of course, the speed of light β^* is given by

$$\beta^* = 1 \text{ meter per meter.}$$

Note also that if a particle has a *constant* speed β , then (2) does yield this exact value from t_1, t_2, x_1 and x_2 .

Next, consider two inertial frames moving relative to each other in such a way that their X -axes are collinear. Call one the laboratory frame and call the second, which moves in a positive direction relative to the first, the rocket frame. A light flashes and is recorded in both systems. The problem is to relate the coordinates (x, y, z, t) in the lab frame to the coordinates (x', y', z', t') in the rocket frame. Under the simplifying assumptions that the flash occurs on the X -axes with $y = z = y' = z' = 0$, and that the origins of the two systems are coincident at $t = 0$, then, if β_r is the constant speed of the rocket frame relative to the lab frame, and if $\beta_r < 1$, the desired relationships are a special case of the well-known Lorentz transformation and are given by

$$(5) \quad x = [x' + \beta_r t'] [1 - \beta_r^2]^{-1/2}$$

$$(6) \quad t = [\beta_r x' + t'] [1 - \beta_r^2]^{-1/2}.$$

With regard to the time of an event, observe that the variable τ , given by

$$(7) \quad \tau = [t^2 - x^2]^{1/2},$$

can be rewritten by means of (3) and (4) as

$$(8) \quad \tau = [(t')^2 - (x')^2]^{1/2}.$$

Since τ is the same in both coordinate frames, it is an invariant which, when $t^2 - x^2 > 0$, is defined to be the proper time of an event. In observing two events, say E_1 with $x = x_1, t = t_1$ and E_2 with $x = x_2, t = t_2$, then

$$(9) \quad \Delta\tau = [(t_2 - t_1)^2 - (x_2 - x_1)^2]^{1/2}$$

is called the proper time between the two events and is also an invariant under transformation (5)–(6).

Finally, let us now turn to the concept of energy. Consider a particle P of mass m which, for simplicity, is in motion only on an X -axis of, say, a lab frame. Its

position is observed at every $\Delta t = (3.335640)10^{-9}$ seconds. Let t_1 and t_2 be the times of two consecutive observations and let x_1 and x_2 be the respective X -coordinates of P at these times. Then the particle's relativistic energy E^* at time t_1 is defined by the forward difference formula

$$(10) \quad E^* = m \frac{t_2 - t_1}{\Delta \tau},$$

where the units of E^* are units of mass. To convert relativistic energy E^* to energy E in conventional units requires ([2], p. 103) multiplication of E^* by c^2 , so that

$$(11) \quad E = E^* c^2.$$

By means of (4), (9), and (10), one can then rewrite (11) as

$$\begin{aligned} E &= m \frac{t_2 - t_1}{\Delta \tau} c^2 \\ &= mc^2 / \left(\frac{\Delta \tau}{t_2 - t_1} \right) \\ &= mc^2 / \left[1 - \left(\frac{x_2 - x_1}{t_2 - t_1} \right)^2 \right]^{1/2} \\ &= mc^2 / (1 - \beta^2)^{1/2}. \end{aligned}$$

If $\beta < 1$, then

$$\begin{aligned} E &= mc^2 \left(1 + \frac{\beta^2}{2} + \frac{3}{8}\beta^4 + \dots \right) \\ &= mc^2 + \frac{m\beta^2 c^2}{2} + \dots \\ &= mc^2 + \frac{mv^2}{2} + \dots. \end{aligned}$$

For β small, then,

$$(12) \quad E \sim mc^2 + \frac{mv^2}{2},$$

where $mv^2/2$ is the kinetic energy of the particle and mc^2 is called its rest energy, because, when $v = 0$,

$$(13) \quad E = mc^2.$$

Thus, the well-known formula (13) has followed directly from difference formulations (2), (3) and (10) of the basic physical concepts of velocity and energy.

It should be noted that, in a consistent fashion, other physical concepts also

can be given relativistic formulations in terms of differences ([2], pp. 103–121). Thus, for example, in the notation used to define E^* in (10), the energy-momentum vector can be defined as $(m(t_2 - t_1)/\Delta\tau, m(x_2 - x_1)/\Delta\tau)$.

References

1. D. Greenspan, Introduction to Numerical Analysis and Applications, Markham, Chicago, 1971.
2. E. F. Taylor and J. A. Wheeler, Spacetime Physics, Freeman, San Francisco, 1966.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

REACHABILITY PROBLEMS IN VECTOR ADDITION SYSTEMS

B. O. NASH, State University of New York at Buffalo

Vector addition systems have arisen in several areas: program schemata, Karp [1]; numeric algorithms, Karp [2]; formal languages, Ginsburg [3], and Nash [4]. The definitions and notations used here follow Karp [1]. Let $N = \{0, 1, 2, \dots\}$. Let $Z = \{\dots, -2, -1, 0, 1, 2, \dots\}$. Let x and y be ordered n -tuples over Z with

$$x = (x_1, x_2, \dots, x_n) \text{ and } y = (y_1, y_2, \dots, y_n),$$

then define $x \leq y$ if $x_i \leq y_i$ for each i , where $i = 1, 2, \dots, n$. Define $x + y$ as $(x_1 + y_1, \dots, x_n + y_n)$ and $-x$ as $(-x_1, -x_2, \dots, -x_n)$. Let $x = (x_1, x_2, \dots, x_m)$ and $y = (y_1, y_2, \dots, y_n)$ be tuples over Z . Define $x \times y$ as the $(m + n)$ -tuple

$$(x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n).$$

Let 0 denote the zero n -tuple where n is clear from context and 0_n if not.

An **n -dimensional vector addition system** is an ordered pair (y, W) where y is an n -tuple over N and W is a finite set of n -tuples over Z .

The **reachability set** $R(y, W)$ of vector addition system (y, W) is the set of all n -tuples x over N , such that either $x = y$ or there exists a sequence w_1, w_2, \dots, w_k of n -tuples, each in W , such that for each i , where $i = 1, 2, \dots, k$, the sum $y + w_1 + w_2 + \dots + w_i \geq 0$ and

$$x = y + w_1 + w_2 + \cdots + w_k.$$

An n -tuple x is **reachable** in (y, W) if x is in $R(y, W)$.

The **general reachability problem** is "Given an n -tuple x and an n -dimensional vector addition system (y, W) , does there exist an algorithm to decide whether or not x is in $R(y, W)$?"

The terms "algorithm" and "decide" are used here in the sense of "computable by Turing Machine," see Rogers [5]. In what follows, the phrase "is x in $R(y, W)$?" will stand for the full version given above.

The general reachability problem is now shown to be equivalent (see Rogers [5]) to two subproblems.

The **reachable-from-zero problem** is "Is x in $R(0, W)$?"

The **zero-reachable problem** is: "Is 0 in $R(y, W)$?"

The reachable-from-zero problem was first studied by M. Rabin [personal communication]. Theorem 1 was discovered by the author; Theorem 2 was discovered independently by M. Rabin and the author after joint discussions of these problems.

The term "reducible" is used as follows: If problem A is reducible to problem B, then, if problem B can be solved, problem A can be solved. If A is reducible to B and B reducible to A, then A and B are equivalent. See Rogers [5].

THEOREM 1. *The general reachability problem is reducible to the zero-reachable problem.*

Proof. Let x be an n -dimensional vector over N and let (y, W) be an n -dimensional vector addition system.

Let $(y \times 1, W')$ be the $(n+1)$ -dimensional vector addition system in which

$$W' = \{w \times 0 \mid w \in W, 0 \in N\} \cup \{(-x) \times (-1)\}.$$

It is claimed that x is in $R(y, W)$ if and only if 0 is in $R(y \times 1, W')$ and therefore that the general reachability problem "is x in $R(y, W)$?" is solved by the zero-reachable problem "is 0 in $R(y \times 1, W')$?" If x is in $R(y, W)$ by using the sequence

$$w_1, w_2, \dots, w_k$$

of vectors from W' then 0 is in $R(y \times 1, W')$ by the sequence

$$w_1 \times 0, w_2 \times 0, \dots, w_k \times 0, (-x) \times (-1)$$

of vectors from W' . If 0 is in $R(y \times 1, W')$ by a sequence of the form

$$w_1 \times 0, w_2 \times 0, \dots, w_k \times 0, (-x) \times (-1),$$

then x is in $R(y, W)$ by the sequence

$$w_1, w_2, \dots, w_k.$$

It only remains to show that if 0 is in $R(y \times 1, W')$ then a sequence ending in $(-x) \times (-1)$ can always be found. But any sequence starting from $y \times 1$ and reaching 0 must contain exactly one occurrence of the vector $(-x) \times (-1)$ to remove the "1" in the $(n+1)$ component introduced by $y \times 1$ since $(-x) \times (-1)$ is the only vector in W' that contains -1 as $(n+1)$ -component. Since $(-x) \times (-1)$ contains no positive components, vectors used after it in the sequence could just as well have been used ahead of it. Therefore it can be moved to the end of the sequence without invalidating the boundary condition and a sequence of the required form can always be found. ■

THEOREM 2. *The zero-reachable problem is reducible to the reachable-from-zero problem.*

Proof. Let x be an n -dimensional vector over N . Let W be a finite set of n -dimensional vectors over z .

Let $W' = \{-w \mid w \in W\}$.

Claim 0 is in $R(x, W)$ if and only if x is in $R(0, W')$.

For any n -dimensional vectors x, y, z , if $x + y = z$ then $z + (-y) = x$. Therefore if the sequence

$$w_1, w_2, \dots, w_k$$

shows that 0 is in $R(x, W)$, then the sequence

$$-w_k, -w_{k-1}, \dots, -w_1$$

shows that x is in $R(0, W')$. The same argument shows that if x is in $R(0, W')$ then 0 is in $R(x, W)$. ■

THEOREM 3. *These problems are equivalent:*

- (a) *the general reachability problem,*
- (b) *the zero-reachable problem,*
- (c) *the reachable-from-zero problem.*

Proof. Theorems 1 and 2 show that (a) is reducible to (b) is reducible to (c). But any reachable-from-zero problem is also a general reachability problem showing that (c) is reducible to (a) and that the three problems are equivalent.

A. Rosenberg [personal communication] has formulated these problems using rational numbers as follows: If any vector $v = (v_1, v_2, \dots, v_k)$ over Z is mapped to the rational number

$$r = p_1^{v_1} p_2^{v_2} \dots p_k^{v_k}$$

in which p_i is the i th prime for $i = 1, 2, \dots, k$, then addition of vectors becomes multiplication of rational numbers. The zero-reachable problem can be restated: Given $s \in N$ and s a finite set of rational numbers, does there exist a sequence

$$s_1, s_2, \dots, s_k$$

of members of s such that for all $i = 1, 2, \dots, k$ the product $s_1 s_2 \cdots s_k$ is in N and $s_1 s_2 \cdots s_k = 1$.

The following results are from Karp [1]:

It is decidable whether $R(y, W)$ is finite.

It is undecidable whether $R(y, W) = R(y', W')$.

From Ginsburg [3] comes:

For all $w \in W$ if $w \geq 0$ then "is x in $R(y, W)$?" is decidable.

In Nash [4], the general reachability problem is shown to be equivalent to the emptiness problem for context-free parallel leveled grammars, a type of formal grammar investigated by the author in his doctoral thesis.

Acknowledgement. The author wishes to acknowledge the financial assistance of the National Research Council of Canada under Grant A-1617 during this work.

References

1. R. M. Karp and R. E. Miller, Parallel program schemata, *J. Comput. System Sci.*, 3 (1969) 147-195.
2. R. M. Karp, R. E. Miller, and S. Winograd, The organization of computations for uniform recurrence equations, *J. Assoc. Comput. Mach.*, 14 (1967) 563-590.
3. S. Ginsburg, *The Mathematical Theory of Context-Free Languages*, McGraw-Hill, New York, 1966.
4. B. O. Nash, Context-Free Parallel Leveled Languages, Research Report CSRR 2026, Dept. Appl. Analysis and Comput. Sci., University of Waterloo, Canada, September, 1970.
5. H. Rogers, *Theory of Recursive Functions and Effective Computability*, McGraw-Hill, New York, 1967.

WHEN DO ALL k -SEQUENCES MODULO m HAVE PERIOD ONE?

E. A. PARBERRY and NANCY GRAUDONS, Wells College

For any pair (k, m) of positive integers we define the set of k -sequences mod m to be the set of sequences $(u_n)_{n \geq 1}$ of integers satisfying $0 \leq u_n < m$, and $u_{n+1} \equiv u_n + [u_n/k] \pmod{m}$, where brackets denote integer part. Obviously, these sequences are ultimately periodic. We are concerned with determining those pairs (k, m) such that every k -sequence mod m has period one.

For example, observe that all 2-sequences mod 12 have period one. We need only consider the sequences

$$(0, 0, \dots)$$

$$(1, 1, \dots)$$

$$(2, 3, 4, 6, 9, 1, 1, \dots)$$

$$(5, 7, 10, 3, 4, 6, 9, 1, 1, \dots)$$

$$(8, 0, 0, \dots)$$

$$(11, 4, 6, 9, 1, 1, \dots),$$

since all other 2-sequences mod 12 are “tails” of these. To see that other periods are possible, note that the 2-sequence mod 10, $(2, 3, 4, 6, 9, \overline{3, 4, 6, 9}, \dots)$, has period four.

Trivially, for $m \leq k$, all k -sequences mod m have period one. If we let $m(k)$ denote the smallest number such that some k -sequence mod $m(k)$ has period larger than one, then we have:

PROPOSITION. *If $k > 1$, then $k(k+1) < m(k) < k(k+2)$.*

Also, letting $i(k)$ be that number such that $m(k) = k(k+1) + i(k)$, we have calculated the following:

$$k: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16$$

$$i(k): 1, 2, 3, 2, 2, 4, 1, 2, 4, 1, 5, 1, 11, 5, 5.$$

These values give no hint toward an exact formula for $m(k)$, but they do show that the bounds given are best possible. To prove the proposition we need the following lemma about the function f , defined by $f(n) = n + [n/k]$ for integers $n \geq 0$.

LEMMA. *The function $f: \mathbb{Z}^+ \rightarrow \mathbb{Z}^+$ is injective, strictly increasing, and its range is*

$$\{n \in \mathbb{Z}^+ : n \not\equiv k \pmod{k+1}\}.$$

Proof. That f is strictly increasing, and hence injective, is obvious. To determine the range, let $n \in \mathbb{Z}^+$ be written $n = c(k+1) + b$, where $0 \leq b \leq k$. If $b \neq k$, then $f(ck+b) = c(k+1) + b$, and hence n is in the range. Also, the consecutive integers $ck+k-1$ and $ck+k$ map by f into $c(k+1)+k-1$ and $c(k+1)+k+1$ respectively, so $c(k+1)+k$ is not in the range of f . ■

Proof of the proposition. The lemma shows that f maps $\{n \in \mathbb{Z}^+ : 0 \leq n \leq m-1\}$ monotonically onto

$$\{n \in \mathbb{Z}^+ : 0 \leq n \leq (m-1) + [(m-1)/k], n \not\equiv k \pmod{k+1}\}.$$

Hence $m \leq f(n) \leq (m-1) + [(m-1)/k]$ implies $0 \leq f(n) - m < k$ just if $m \leq k(k+1)$. Therefore every k -sequence mod m has ultimate period one if $m \leq k(k+1)$, so $m(k) > k(k+1)$.

Denoting the r th iterate of f on n by $f^r(n)$, we must have $f^i(k) = k(k+1) + j$ with $0 \leq j < k$ for some unique i , since $f(n) > n$ when $n \geq k$, and if $f^{i-1}(k) < k(k+1) \leq f^i(k)$, then $f^i(k) \leq f(k^2+k-1) = k(k+1) + k-1$. Indeed since $i \geq 2$ for $k \geq 2$, the lemma shows $f^{i-1}(k) < k^2+k-1$, whence $f^i(k) \leq k(k+1) + k-2$ and $0 \leq j \leq k-2$. Now, letting $m = k(k+1) + j + 1$, we have

$f^{i+1}(k) = k(k+1) + j + k + 1 \equiv k \pmod{m}$, so $f^{i+1}(k) \equiv f^0(k) \pmod{m}$. Therefore, the k -sequence $\text{mod}(k(k+1) + j + 1)$ whose initial term is k has period $i+1$. With $k \geq 2$, we have $i \geq 2$ and $0 \leq j \leq k-2$ so

$$m(k) \leq k(k+1) + j + 1 < k(k+2). \quad \blacksquare$$

We have not been able to make progress on the following problem. Let $M(k)$ denote the largest integer such that all k -sequences $\text{mod } M(k)$ have ultimate period one. Does $M(k)$ exist for any $k > 1$; and if so, how large is it? Preliminary calculations show that if $M(2)$ exists, then $M(2) \geq 197$.

CLASSROOM NOTES

EDITED BY ROBERT GILMER

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306; notes are usually limited to three printed pages.

ON INJECTIVE MODULES

AZMI HANNA, American University of Beirut

Let R be a ring, not necessarily with an identity element. Among the several characterizations of injective R -modules, one finds the following:

An R -module Q is injective if and only if every monomorphism with domain Q has a left inverse.

The necessity of the condition is immediate, while the proof of the sufficiency is usually given after showing that every R -module can be embedded in an injective module. We give an elementary proof of the sufficiency of the condition which avoids use of this embedding. An analogous proof holds for the dual theorem that an R -module P is projective if and only if every epimorphism with codomain P has a right inverse.

Let Q be an R -module such that every monomorphism with domain Q has a left inverse. Let $u: E \rightarrow F$ be an R -module monomorphism, $f: E \rightarrow Q$ a homomorphism. We shall show the existence of a homomorphism $g: F \rightarrow Q$ such that $gu = f$. To do so, form the pushout diagram

$$\begin{array}{ccc} E & \xrightarrow{u} & F \\ f \downarrow & & \downarrow f' \\ Q & \xrightarrow{u'} & P \end{array}$$

of the homomorphism u and f . Here $P = (F \oplus Q)/N$, where N is the submodule of $F \oplus Q$ consisting of all pairs $(u(a), -f(a))$ with $a \in E$. $f'(X) = N + (x, 0)$, $u'(y) = N + (0, y)$ for every $x \in F$ and $y \in Q$. The above diagram is commutative: $f'u = u'f$. Further, u' is a monomorphism, for if $u'(y) = N + (0, y) = 0$, then $(0, y) = (u(a), -f(a))$ for some $a \in E$. Since u is a monomorphism, we must have $a = 0$ and consequently $y = 0$. By assumption, there exists a homomorphism $v: P \rightarrow Q$ such that $vu' = 1_Q$. Let $g = vf': F \rightarrow Q$. Then $gu = vf'u = vu'f = f$. Thus Q is injective.

The above proof and its dual can be generalized to any abelian category which may fail to have enough injectives or enough projectives. For instance, it is shown in [1, p.260] and in [2, p.257] that the category of sheaves over a topological space has enough injectives but not necessarily enough projectives.

References

1. R. Godement, *Topologie algébrique et théorie des faisceaux*, Hermann, Paris, 1958.
2. B. Mitchell, *Theory of Categories*, Academic Press, New York, 1965.

THE HAMEL DIMENSION OF ANY INFINITE DIMENSIONAL SEPARABLE BANACH SPACE IS c .

H. ELTON LACEY, University of Texas at Austin

In this MONTHLY [1] an elementary proof of the fact that an infinite dimensional Banach space cannot have dimension \aleph_0 is presented. A simple argument can also establish (without the continuum hypothesis) the statement in the title of this note.

Let X be an infinite dimensional Banach space. It suffices to demonstrate a linearly independent set in X of cardinality c . By the Hahn-Banach theorem and the fact that X is infinite dimensional, there are sequences $\{x_n\}$ in X and $\{x_n^*\}$ in X^* such that $x_n^*(x_n) \neq 0$ and $x_n^*(x_m) = 0$ for $n \neq m$. In particular, $\{x_n\}$ is linearly independent and for each n , x_n is not in the closed linear span of $\{x_m: m \neq n\}$. Now let $\{N_t\}_{0 < t < 1}$ be a family of subsets of the positive integers such that $N_t \cap N_{t'}$ is finite for $t \neq t'$ and each N_t is infinite (see [2] for an easy proof of this). It is easy to see that the family $\{x_t\}_{0 < t < 1}$, where $x_t = \sum_{n \in N_t} x_n / 2^n$ is a linearly independent family in X .

If, in addition, X is separable, it follows that the Hamel dimension of X is c since the cardinality of X is c .

For a recent result on dimension, see [3] where the authors show that the only barreled spaces of countable Hamel dimension are those isomorphic to the space of finitely non-zero sequences in its strongest locally convex topology.

References

1. W. R. Bauer, and R. H. Benner, The non-existence of a Banach space of countably infinite Hamel dimension, this MONTHLY, 78 (1971) 895-896.
2. J. R. Buddenhagen, Subsets of a countable set, this MONTHLY, 78 (1971) 536-537.
3. S. Saxon, and M. Levin, Codimensional subspaces of a barreled space, Proc. Amer. Math. Soc., 29 (1971) 91-96.

A NOTE ON CONFORMALITY

R. K. WILLIAMS, Southern Methodist University

In the following note, it is shown how a restricted orthogonality requirement implies analyticity. As a special case, we get the well-known result that conformality implies analyticity.

Suppose that $f(z) = u + iv$ is a transformation of a domain D in the z -plane into the z -plane. Let u and v be continuous with continuous partial derivatives in D , and let the Jacobian of the transformation

$$J = \begin{vmatrix} u_x & u_y \\ v_x & v_y \end{vmatrix},$$

be nonzero in D . (Hence $f(z)$ is locally one-to-one in D .)

For $z_0 \in D$, let $L_1(z_0)$, $L_2(z_0)$, $L_3(z_0)$, and $L_4(z_0)$ be line segments in D , through z_0 , and having angles of inclination 0 , $\pi/2$, $\pi/4$, and $3\pi/4$ respectively. Suppose that for each $z_0 \in D$, $f(L_1(z_0))$ and $f(L_3(z_0))$ are orthogonal to $f(L_2(z_0))$ and $f(L_4(z_0))$ at $f(z_0)$, respectively. We then have the following:

THEOREM. *Either $f(z)$ or $\overline{f(z)}$ is analytic in D .*

Proof. Let $z_0 = x_0 + iy_0 \in D$. Then parametric representations for $L_1(z_0)$, $L_2(z_0)$, $L_3(z_0)$, and $L_4(z_0)$ are

$$\begin{cases} x = x_0 + t, \\ y = y_0, \end{cases} \quad \begin{cases} x = x_0, \\ y = y_0 + t, \end{cases} \quad \begin{cases} x = x_0 + t, \\ y = y_0 + t, \end{cases} \quad \begin{cases} x = x_0 + t, \\ y = y_0 - t. \end{cases}$$

Also, parametric representations for $f(L_1(z_0))$, $f(L_2(z_0))$, $f(L_3(z_0))$, and $f(L_4(z_0))$ are

$$\begin{cases} u = u(x_0 + t, y_0), \\ v = v(x_0 + t, y_0), \end{cases} \text{ etc.}$$

Letting T_1 , T_2 , T_3 , and T_4 be tangent vectors to $f(L_1(z_0))$, $f(L_2(z_0))$, $f(L_3(z_0))$, and $f(L_4(z_0))$ at $f(z_0)$, and using complex notation, we see that

$$T_1 = u_x + iv_x, \quad T_2 = u_y + iv_y, \quad T_3 = (u_x + u_y) + i(v_x + v_y), \quad \text{and}$$

$$T_4 = (u_x - u_y) + i(v_x - v_y),$$

where the partial derivatives are evaluated at z_0 . Since by assumption, T_1 and T_3 are orthogonal to T_2 and T_4 respectively, we have

$$(1) \quad u_x u_y + v_x v_y = 0$$

$$(2) \quad u_x^2 - u_y^2 + v_x^2 - v_y^2 = 0.$$

From (1) we see that there is a number $c = c(z_0)$ such that

$$(3) \quad u_y = -cv_x \text{ and } v_y = cu_x.$$

Using (2) and (3), we have

$$u_x^2 + v_x^2 = u_y^2 + v_y^2 = c^2(v_x^2 + u_x^2).$$

Since $J \neq 0$, $u_x^2 + v_x^2 \neq 0$, so $c(z_0) = \pm 1$ for each $z_0 \in D$. Equations (3) imply that $c(z)$ is continuous at each $z_0 \in D$. Thus $c(z) \equiv 1$ or $c(z) \equiv -1$.

If $c(z) \equiv 1$, equations (3) imply that $f(z)$ is analytic in D . If $c(z) \equiv -1$, the analyticity of $\overline{f(z)}$ follows similarly.

COROLLARY. *Under the hypotheses of the theorem, $f(z)$ is either angle preserving or angle reversing in D .*

Proof. From the theorem and the fact that $J \neq 0$ in D , either $f(z)$ or $\overline{f(z)}$ is analytic with a nonvanishing derivative in D . Thus either $f(z)$ or $\overline{f(z)}$ is angle preserving in D . The corollary follows.

It is clear that conformality implies our (seemingly weaker) orthogonality condition. The well-known fact that conformality implies analyticity now follows from the proof of the corollary.

A WRONSKIAN CONDITION RELATED TO ORDINARY DIFFERENTIAL EQUATIONS

L. C. EGGAN and A. J. INSEL, Illinois State University

It is well known that in general the vanishing of the Wronskian of $n + 1$ functions on the real line \mathbf{R} is not sufficient to establish their linear dependence on \mathbf{R} . In a differential equations course recently one of us asked the question:

Under what conditions does the Wronskian of a finite sequence of functions being zero imply that the functions are linearly dependent?

Naturally the expected answer was:

It is sufficient that the functions all be solutions to a differential equation of the form

$$a_m y^{(m)} + a_{m-1} y^{(m-1)} + \cdots + a_1 y' + a_0 y = 0,$$

where $a_m, a_{m-1}, \dots, a_1, a_0$ are continuous on an interval over which a_m is never zero.

One student suggested (actually stated) that it is sufficient for the sequence to consist of a function and its derivatives.

In this note we prove this assertion and as a corollary obtain a sufficient condition for a function to be analytic. Our proof is a refinement of the following theorem to be found in [1, p. 47] which is stated here for completeness.

THEOREM 8. Let $\phi_1(t), \phi_2(t), \dots, \phi_n(t)$ be any n functions continuous together with their first $(n-1)$ derivatives over some interval $t_1 < t < t_2$.

(a) If the ϕ 's are linearly dependent over $t_1 < t < t_2$, then their Wronskian

$$W(\phi_1, \phi_2, \dots, \phi_n) \equiv 0, \quad t_1 < t < t_2.$$

(b) Suppose

(1) $W(\phi_1, \phi_2, \dots, \phi_n) \equiv 0, \quad t_1 < t < t_2$; but

(2) for some $(n-1)$ of the ϕ 's (say, without loss of generality, all but ϕ_n) $W(\phi_1, \phi_2, \dots, \phi_{n-1}) \neq 0$, all $t, t_1 < t < t_2$, then $\phi_1, \phi_2, \dots, \phi_n$ are linearly dependent over $t_1 < t < t_2$.

With the aid of the above theorem and a modest background in real variable theory our theorem can be presented in a first year differential equations course. The converse of our theorem is well known (cf. [2], pp. 99–100).

THEOREM. If f is a real valued function defined on the real line \mathbf{R} and if there exists a positive integer n such that f is $2n$ -fold differentiable on \mathbf{R} and $W(f, f', \dots, f^{(n)}) \equiv 0$ on \mathbf{R} , then f is the solution of an n -th order homogeneous linear differential equation with constant coefficients not all of which are zero. In particular, f is analytic.

Proof. We first suppose n is minimal in this regard, so

$$(1) \quad W(f, f', \dots, f^{(n-1)}) \neq 0$$

at some point of \mathbf{R} . Hence by the continuity of the first $2(n-1)$ derivatives of f , (1) holds on some connected set I in \mathbf{R} . Choose I to be a maximal connected set on which (1) holds, and note, again by continuity, that I is open.

Now apply Theorem 8(b) as stated above to conclude that $f, f', \dots, f^{(n)}$ are linearly dependent on I . (Note that, in the context of this paper, it is required in the statement of Theorem 8 that $f^{(2n)}$ be continuous. However, the proof makes no use of this hypothesis and hence the continuity of $f^{(2n)}$ may be ignored.) Continuing with the argument, by linear dependence there exist real numbers a_0, a_1, \dots, a_n , not all zero, such that $\sum_{j=0}^n a_j f^{(j)} = 0$ on I . Let g be the solution to the differential equation

$$(2) \quad \sum_{j=0}^n a_j y^{(j)} = 0$$

which is defined on all of \mathbf{R} and is such that $f = g$ on I . Then $f^{(j)} = g^{(j)}$ on I and hence, by continuity, also on the closure of I , which we denote by \bar{I} , for $0 \leq j \leq 2(n-1)$. Thus we must have

$$(3) \quad W(f, f', \dots, f^{(n-1)}) = W(g, g', \dots, g^{(n-1)}) \text{ on } \bar{I}.$$

Now since $g, g', \dots, g^{(n-1)}$ are solutions to (2) whose Wronskian does not vanish on I , it is well known (cf. [1], Theorem 7, p. 47) $W(g, g', \dots, g^{(n-1)})$ does not vanish on all of \mathbf{R} . Thus by (3) inequality (1) holds on I . But I was maximal, so the open set I must equal \bar{I} . This can happen only if $I = \mathbf{R}$, and therefore $f = g$ on \mathbf{R} as desired.

References

1. W. Hurewicz, Lectures on Ordinary Differential Equations, Technology Press of M. I. T. and Wiley, New York, 1958.
2. E. Rainville, Elementary Differential Equations, 3rd edition, Macmillan, New York, 1964.

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to either of the editors: J. G. Harvey, Department of Mathematics, University of Wisconsin, WI53706; M. W. Pownall, Department of Mathematics, Colgate University, Hamilton, NY 13346.

TEACHING APPLICABLE MATHEMATICS

E. A. BENDER, Institute for Defense Analyses

Introduction. In this article I shall examine what appear to be the purposes of the usual two year calculus service course, (I include differential equations in "calculus") what I think they should be, and how these latter could be achieved. What I have to say is not intended for courses for undergraduates planning to enter pure mathematics, but those students might also profit from the proposed approach.

I want to suggest a philosophy — an approach. As a result the details of carrying out the ideas are not described. Filling in these details amounts to writing a course which is by no means trivial. Although it would be difficult, I believe it is not impossible.

Purposes of the course. There seem to be three traditional goals for the calculus sequence:

- (i) Learn some useful mathematics.
- (ii) Develop a feel for one or more areas of mathematics. (Analytic geometry, linear algebra or elementary probability are sometimes included.)
- (iii) Obtain an idea of what mathematicians do and of the beauty of mathematics.

We regard (i) as the analytic version and (ii) as the synthetic version of the same idea. The last goal might be labeled "culture" and is frequently absent. To some extent (ii) and (iii) are reasonable outgrowths of (i).

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

ASSOCIATE EDITORS: JOSHUA BARLAZ, ERIC S. LANGFORD. COLLABORATING EDITORS: LEONARD CARLITZ, GULBANK D. CHAKERIAN, HASKELL COHEN, S. ASHBY FOOTE, ISRAEL N. HERSTEIN, MURRAY S. KLAMKIN, DANIEL J. KLEITMAN, ROGER C. LYNDON, MARVIN MARCUS, CHRISTOPH NEUGEBAUER, ALBERT WILANSKY, and UNIVERSITY OF MAINE PROBLEMS GROUP: EARL M. L. BEARD, GEORGE S. CUNNINGHAM, CLAYTON W. DODGE, OSKAR FEICHTINGER, WILLIAM R. GEIGER, GARY HAGGARD, PHILIP M. LOCKE, JOHN C. MAIRHUBER, CURTIS S. MORSE, GRATTAN P. MURPHY, EDWARD S. NORTHAM and WILLIAM L. SOULE, JR.

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before June 30, 1973. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.

An asterisk () means neither the proposer nor the editors supplied a solution.*

E 2403. Proposed by W. A. Al-Salam, University of Alberta, and A. M. Chak, West Virginia University

It is known that a generalization of the binomial theorem is Euler's identity

$$(1) \quad (1+x)(1+qx)\cdots(1+q^{n-1}x) = \sum_{k=0}^n \frac{F(n,q)q^{k(k-1)/2}}{F(n-k,q)F(k,q)} x^k,$$

where q is a fixed complex number which is not a primitive k th root of unity for any $k \geq 2$, where

$$F(k,q) = \prod_{j=1}^k (1+q+\cdots+q^{j-1}),$$

and where $F(0,q) = 1$. (Thus $F(k,1) = k!$.)

Show that the only solution of

$$(2) \quad (1+x)(1+c_1x)\cdots(1+c_{n-1}x) = \sum_{k=0}^n \frac{\phi_n}{\alpha_{n-k}\beta_k} x^k$$

is this identity of Euler. That is, if given a sequence c_1, c_2, \dots of complex numbers there exist sequences $\phi_0, \phi_1, \dots; \alpha_0, \alpha_1, \dots; \beta_0, \beta_1, \dots$ such that (2) holds identically for all values of n , then necessarily $c_1 = c_n/c_{n-1} = q$ is constant for $n = 2, 3, \dots$ and thus (2) must have the form of (1).

E 2404*. *Proposed by Russell Maurer, Harvard Medical School*

At Smith College, the graduation exercises traditionally proceed as follows: Although each diploma is made out to a particular girl, all the diplomas are initially given out at random. All of the girls who do not get their own diplomas then form a circle, and each passes the diploma she has to the girl on her right. Those who now have their own diplomas drop out, and the remaining girls again pass their diplomas to the right, and so on. This procedure is repeated until each girl has her own diploma. If there are n girls in the graduating class, what is the probability that it takes precisely k passes before each girl has her own diploma?

E 2405. *Proposed by R. E. Shafer, Lawrence Radiation Laboratory*

Forman S. Acton in his book *Numerical Methods that (almost) Work* (Harper-Row, New York, 1970, pp. 29–40), proposed several numerical methods for the evaluation of

$$F(b) = \int_0^\infty \frac{\tan^{-1} bx}{1+x^2} dx.$$

Derive the following additional form:

$$F(b) = -\frac{1}{2} \log b \log \frac{1+b}{1-b} + \sum_{n=0}^{\infty} \frac{b^{2n+1}}{(2n+1)^2}, \quad 0 < b < 1.$$

E 2406. *Proposed by Erwin Just and Norman Schaumberger, Bronx Community College*

What is the maximum value of α and the minimum value of β for which

$$\left(1 + \frac{1}{n}\right)^{n+\alpha} \leq e \leq \left(1 + \frac{1}{n}\right)^{n+\beta}$$

for all positive integers n ?

E 2407. *Proposed by A. W. Walker, Toronto, Canada*

Given the circumcenter O , orthocenter H , and incenter I of an unknown triangle T , (A) locate by Euclidean construction the Gergonne point and the Lemoine point of T (i.e., the centers of perspective of T with the triangles formed respectively by the contact points of the sides of T with its incircle and by the tangent lines at the vertices of T to its circumcircle). (B) Locate the orthocenters of the pedal triangles of H and I .

SOLUTIONS OF ELEMENTARY PROBLEMS

A Set of Unrelated Primes

E 2293 [1971, 405; 1972, 302]. *Proposed by Erwin Just, Bronx Community College*

Does there exist an infinite set of primes, S , such that whenever $p \in S$ and $q \in S$ we have $(\frac{1}{2}(p-1), \frac{1}{2}(q-1)) = 1$, $(p, q-1) = 1$ and $(p-1, q) = 1$?

Solution by Frederick Carty, Parsippany, New Jersey. A set with the desired properties does exist. We shall demonstrate this by constructing $S = \{p_1, p_2, \dots\}$ inductively. Let $p_1 = 3$ and $p_2 = 5$, and suppose that p_1, p_2, \dots, p_n have been chosen. Let $d_n = \prod_{i=1}^n \frac{1}{2} p_i (p_i - 1)$; by Dirichlet's theorem, there exists a prime p_{n+1} of the form $2kd_n - 1$. Then for $i \leq n$, obviously $(p_i, p_{n+1} - 1) = (p_i - 1, p_{n+1}) = (\frac{1}{2}(p_i - 1), \frac{1}{2}(p_{n+1} - 1)) = 1$ and we are done.

Also solved by Anders Bager (Denmark), Problem Solving Group Berne (Switzerland), R. T. Bumby, Frederick Carty, M. S. Demos, R. J. Dickson, Harold Donnelly, Neal Felsinger, Heiko Harborth (Germany), C. V. Heuer & G. A. Heuer, Emmett Keeler & Joel Spencer, Harry Lass, L. E. Mattics, Kenneth Schilling, Paul Smith, Karl Stoop (Colombia), Charles Wexler, Gregory Wulczyn, and the proposer. Partial solutions by John Coolidge, and by Bernardo Recamán (Colombia).

Rearrangement of Series

E 2343 [1972, 303]. *Proposed by G. A. Heuer, Concordia College*

According to a well-known theorem of analysis, a series of real numbers is unconditionally convergent (i.e., $\sum a_{\phi(n)} = \sum a_n$ for every permutation ϕ of the positive integers) if and only if it is absolutely convergent. Certain kinds of rearrangements, however, will leave the sum of an arbitrary convergent series unaltered. (A) Prove that if $\phi(n) - n$ is bounded, and $\sum a_n$ is any convergent series, then $\sum a_{\phi(n)} = \sum a_n$. (B) Prove or disprove: If $\phi(n) - n$ is unbounded, then there is a series $\sum a_n$ for which $\sum a_{\phi(n)} \neq \sum a_n$.

Solution by David Monk, Mathematical Institute, Edinburgh, Scotland. To show part (A), let s_m and t_m be the m th partial sums of $\sum a_n$ and $\sum a_{\phi(n)}$ respectively. Suppose that $n - k \leq \phi(n) \leq n + k$ for all n . If $m > k$ and $0 < r \leq m - k$, then $\phi^{-1}(r) - k \leq \phi(\phi^{-1}(r)) = r \leq m - k$ so that $\phi^{-1}(r) \leq m$. Thus

$$\{1, \dots, m - k\} \subseteq \phi(\{1, \dots, m\})$$

so that t_m includes all terms a_n for which $n \leq m - k$. On the other hand, $\phi(n) \leq m + k$ for $n = 1, \dots, m$ so t_m excludes all terms a_n for which $n > m + k$. Therefore whenever $m > k$,

$$|t_m - s_m| \leq |a_{m-k+1}| + \dots + |a_{m+k}|$$

which approaches 0 as $m \rightarrow \infty$ since $a_n \rightarrow 0$ and k is fixed.

As for part (B), the assertion is false. Let ϕ be the permutation which interchanges 1 and 2, 3 and 5, 6 and 9, ..., $\frac{1}{2}n(n+1)$ and $\frac{1}{2}n(n+3)$, ... and leaves all other integers unchanged. Since $\frac{1}{2}n(n+3) - \frac{1}{2}n(n+1) = n$, it follows that $\phi(n) - n$ is unbounded. Suppose that $\sum a_n$ is any convergent series with sum s . We have (in the notation of part (A)) two possibilities for $t_m - s_m$. If $m = \frac{1}{2}n(n+3)$ for some natural number n , then $t_m - s_m = 0$; otherwise $\frac{1}{2}n(n+1) \leq m < \frac{1}{2}n(n+3)$ for some (unique) natural number n and in this case

$$t_m - s_m = a_{n(n+3)/2} - a_{n(n+1)/2}.$$

We conclude that since $s_m \rightarrow s$ likewise $t_m \rightarrow s$.

Also solved by Anders Bager (Denmark), Roby Ballard & George Zahn, J.A. Belward (Australia), M. T. Bird, D. M. Bloom, D. L. Costa, R. J. Dickson, David Farnsworth, Iowa Problem Group, J. R. Kuttler, Joel Levy, H. Sarbadhikari (India), Kenneth Schilling, Gary Thomas, and the proposer.

Editor's comment. Monk also refers to R. P. Agnew, *Permutations preserving convergence of series*, Proc. Amer. Math. Soc. 6 (1955), 563–564, where it is shown that a necessary and sufficient condition that $\sum a_{\phi(n)}$ converge whenever $\sum a_n$ does (and to the same sum) is the existence of an integer N such that for each n the set $\{\phi(r) : 1 \leq r \leq n\}$ is expressible as a union of not more than N blocks of consecutive integers. In the example used in (B) above, $N = 3$.

Jumping Around in an Ellipse

E2345 [1972, 303]. *Proposed by E. S. Langford, University of Maine*

Let S be a nonempty compact subset of the plane. A sequence $\{P_n\}$ of points of S has the following property:

$$d(P_n, P_{n+1}) = \max\{d(P_n, P) : P \in S\}.$$

Let $d_n = d(P_n, P_{n+1})$. Then obviously $d_1 \leq d_2 \leq \dots \leq \delta$, where δ is the diameter of S . Let $d = \lim d_n$. (a) Is it possible that $d < \delta$? (b) Is it possible that the sequence $\{d_n\}$ is strictly increasing? (c) Is it possible that $\{d_n\}$ is strictly increasing and, in addition, that $d < \delta$?

Solution by H. S. Witsenhausen, Bell Telephone Laboratories. The answer to all three questions is yes. In fact, we can show that examples of (a) exist if and only if $\delta \leq \sqrt{3}d$, and that an example for (c) exists for every d such that $\sqrt{3}d > \delta$.

Let $\{P_n\}$ be any sequence of the type described in the problem. For $\varepsilon > 0$, choose n so that $d - \varepsilon \leq d_n$. If $Q \in S$, then $d(Q, P_n) \leq d_n \leq d$ and $d(Q, P_{n+1}) \leq d_{n+1} \leq d$, so that S is contained in the intersection of the two circular disks of radius d centered at P_n and P_{n+1} . The diameter of this intersection does not exceed $\sqrt{3}(d + \varepsilon)$ from which it follows that $\delta \leq \sqrt{3}d$, i.e., that $d \geq \sqrt{3}\delta/3$. We can achieve equality by taking S to be a rhombus which is the union of two equilateral triangles, and then taking for P_n alternately the ends of the shorter diagonal.

One can construct an example with $\{d_n\}$ strictly increasing and $\delta = (\sqrt{3} - \varepsilon)d$ for arbitrary $\varepsilon > 0$ as follows: Let S be the union of an eccentric ellipse (with unit major axis) together with two points symmetrically located on the extensions of the minor axis which are a distance $\sqrt{3} - \varepsilon$ apart. If P_1 is chosen on the ellipse, sufficiently close to one end of the major axis, the sequence $\{P_n\}$ is uniquely determined, the d_n are strictly increasing, and $d = 1$. [This is probably most easily seen by noting that given P_1 , the point P_2 must lie somewhere strictly between the other end of the major axis and the point antipodal to P_1 . This can be shown by standard calculus techniques.—Ed.]

Also solved by J. C. Binz and the Problem Solving Group Berne (Switzerland), J. W. Boyd, R. J. Dickson, D. P. Giesy, G. A. Heuer, Ralph Jones, E. Keeler & J. Spencer, C. J. Knight, O. P. Lossers (Netherlands), Bill Margolis, L. E. Mattics, G. L. Miller, Bill Sands, and David Weinberger.

Editor's Comment: The ellipse plus two points example was used by several solvers. Others used two non-overlapping circles together with two isolated points, and still others used appropriate unions of circular arcs. Dickson and Weinberger (independently) observe that we can take the convex hull in any of the examples above, and still have an example. Jones notes that the only way we can have $\delta = d\sqrt{3}$ is for $\{P_n\}$ to alternate (after the first term) between two points so that d_n is constant for $n \geq 2$.

Groups with Small Centralizers

E 2346 [1972, 303]. *Proposed by Louis Shapiro, Howard University*

Say that a group has *small centralizers* if every non-identity element commutes only with its inverse, itself, and the identity. Characterize all groups with small centralizers.

I. Solution by the Problem Solving Group Berne, Switzerland. We shall show that the only groups with small centralizers are the cyclic groups of orders one, two, and three and the symmetric group S_3 . If $o(G) \leq 3$, then certainly G has small centralizers, so assume that G has small centralizers and that $o(G) \geq 4$. Since $aa^2 = a^2a$ by the associative law, every $g \in G$ other than the identity e has order 2 or 3. But not every element in G other than e can have order 2, for if a and b are distinct elements of order 2 and ab is also, then $ab = (ab)^{-1} = b^{-1}a^{-1} = ba$ contrary to assumption. Hence G must contain at least one element b of order 3; this implies the existence of at least two elements of order 3 since $b^2 \neq b$ is also of order 3.

We show now that G must have exactly two elements of order 3. Suppose to the contrary that $c \in G$ is an element of order 3 distinct from b and b^2 . Then there are at least four elements of order 3: b, b^2, c, c^2 . Consider the element bc . Either $o(bc) = 2$ or $o(bc) = 3$. Suppose that $o(bc) = 2$. Now $cb \neq bc$ by assumption and it is easy to show that $o(cb) = 2$. But $(bc)(cb)$ is not the identity and $(bccb)^2 = bc^2b^2c^2b = bc^2b^{-1}c^{-1}b = bc^2(cb)^{-1}b = bc^2(cb)b = e$. This implies that $o(bccb) = 2$ and therefore that bc and cb commute as above. This is a contradiction.

Suppose that $o(bc) = 3$. Consider the elements bc^2 and b^2c . They are distinct, neither is the identity, and they are not inverses of each other since $(bc^2)^{-1} = c^{-2}b^{-1} = cb^2$ and c and b^2 do not commute by assumption. But $(bc)^2 = (bc)^{-1} = c^{-1}b^{-1} = c^2b^2$ so that $(bc^2)(b^2c) = b(c^2b^2)c = b(bc)^2c = (b^2c)(bc^2)$ contrary to assumption. We have now shown that G has precisely two elements of order 3: b and b^2 .

Since $o(G) \geq 4$, it follows that G has an element g of order 2. Then bg and b^2g are two further elements of G which must be of order 2 and $K = \{e, b, b^2, g, bg, b^2g\}$ is a subgroup of G isomorphic to the symmetric group S_3 . This must exhaust G for suppose there exists an element h not in K . Necessarily $o(h) = 2$ and since $gh \notin K$ it follows that $o(gh) = 2$ and therefore that $gh = hg$ as before. Since S_3 obviously has small centralizers, we see that G has small centralizers if and only if it has order not exceeding 3 or is isomorphic to S_3 , that is, if and only if G is isomorphic to a subgroup of S_3 .

II. *Solution (G finite) by D. M. Bloom, Brooklyn College.* If $o(G) = n$ and there are a and b conjugacy classes of elements of orders 2 and 3 respectively, then since each element centralizes only its own powers, we have for the class equation of G

$$a \frac{n}{2} + b \frac{n}{3} + 1 = n;$$

that is, $n(6 - 3a - 2b) = 6$ so that n is a divisor of 6. If $n = 1, 2$, or 3 we have the cyclic groups which do have small centralizers, whereas if $n = 6$, G cannot be abelian, and must therefore be S_3 which does have small centralizers.

III. *Solution (G finite) by H. D'Alarcao and T. Moore, Bridgewater State College.* If G is finite and non-trivial and has small centralizers, then so does every subgroup of G , in particular every Sylow p -subgroup G_p of G . Since G_p has non-trivial center, $o(G_p) = p$. But every element of G has order either 2 or 3 so that every Sylow p -subgroup of G has order either 2 or 3. Hence the order of G is either 1, 2, 3, or 6. The cyclic group of order 6 does not have small centralizers, so that the only groups with small centralizers are the cyclic groups of order 1, 2, or 3 and the symmetric group S_3 .

IV. *Solution (infinite case) by John Comiskey, Monsignor Farrell High School (New York City).* Having determined the finite groups with small centralizers, suppose that G is an infinite group with small centralizers. Choose seven elements of G and consider the subgroup H generated by these elements. A group generated by a finite number of elements whose orders do not exceed 3 is finite (see B. H. Neumann, *Groups whose elements have bounded orders*, J. London Math. Soc. 11 (1937), 195–198) so that H is a finite group with small centralizers whose order exceeds 6. This is a contradiction so that there does not exist an infinite group with small centralizers.

Complete solutions submitted by Anders Bager (Denmark), the Bennett College Team, D. M. Bloom, D. E. Bridgewater, John Comiskey, D. A. Sibley, Brian Wesselink, and the proposer. Partial solutions (assuming G finite) by H. D'Alarcao and T. Moore, Annette Dittmer, M. G. Greening (Australia), C. V. Heuer, Barbara Keller, Desmond MacHale (England), David Ritter, Steven Russ, Bruce Staal, and S. Srinivasan (India). Partial solution by M. R. Modak (India).

Editor's comment. For the case of G infinite, most solvers made reference to the special case $n = 3$ of the Burnside problem to reduce the problem back to the finite case. (See Marshall Hall, *Theory of Groups*, Macmillan, 1959, Section 18.2.) The solution by the Berne Group is interesting because it does not use this result.

Symmedian Point of a Triangle

E 2347 [1972, 303]. *Proposed by Leonard Carlitz, Duke University*

Let P denote a point in the interior of the triangle ABC . Let α, β, γ denote the angles of ABC . Let R_1, R_2, R_3 denote the distances from P to the vertices of ABC , and let r_1, r_2, r_3 denote the distances from the sides of ABC . Show that

$$R_1^2 \sin^2 \alpha + R_2^2 \sin^2 \beta + R_3^2 \sin^2 \gamma \leq 3(r_1^2 + r_2^2 + r_3^2)$$

with equality if and only if P is the symmedian point of ABC .

Solution by Ralph Garfield, The College of Insurance, New York. Let $\theta_1 = \angle PAC$ and $\theta_2 = \angle BAP$. We then see that $r_1 = R_1 \sin \theta_2$ and $r_2 = R_1 \sin \theta_1$, so that

$$\begin{aligned} r_2 &= R_1 \sin(\alpha - \theta_2) = R_1(\sin \alpha \cos \theta_2 - \cos \alpha \sin \theta_2) \\ &= R_1 \sin \alpha \cos \theta_2 - r_1 \cos \alpha. \end{aligned}$$

Therefore

$$r_1^2 \sin^2 \alpha + (r_2 + r_1 \cos \alpha)^2 = R_1^2 \sin^2 \theta_2 \sin^2 \alpha + R_1^2 \sin^2 \alpha \cos^2 \theta_2$$

which is $r_1^2 + r_2^2 + 2r_1 r_2 \cos \alpha = R_1^2 \sin^2 \alpha$. It now follows that

$$\begin{aligned} &R_1^2 \sin^2 \alpha + R_2^2 \sin^2 \beta + R_3^2 \sin^2 \gamma \\ &= 2(r_1^2 + r_2^2 + r_3^2) + 2(r_1 r_2 \cos \alpha + r_1 r_3 \cos \beta + r_2 r_3 \cos \gamma). \end{aligned}$$

To complete the problem it suffices to show that

$$2(r_1 r_2 \cos \alpha + r_1 r_3 \cos \beta + r_2 r_3 \cos \gamma) \leq r_1^2 + r_2^2 + r_3^2.$$

Using Lagrange multipliers, we maximize

$$L(\alpha, \beta, \gamma, \lambda) = 2(r_1 r_2 \cos \alpha + r_1 r_3 \cos \beta + r_2 r_3 \cos \gamma) + \lambda(\alpha + \beta + \gamma - \pi).$$

We find that

$$\lambda = 2r_1 r_2 \sin \alpha = 2r_1 r_3 \sin \beta = 2r_2 r_3 \sin \gamma.$$

Since none of r_1, r_2, r_3 is equal to zero, we have

$$\begin{aligned} r_2 \sin \alpha &= r_3 \sin \beta, \\ r_1 \sin \alpha &= r_3 \sin \gamma = r_3 \sin(\alpha + \beta) = r_3 (\sin \alpha \cos \beta + \cos \alpha \sin \beta) \\ &= r_3 \sin \alpha \cos \beta + r_2 \sin \alpha \cos \alpha. \end{aligned}$$

Now assuming $\sin \alpha \neq 0$ we find

$$(r_1 - r_2 \cos \alpha)^2 + (r_1 \sin \alpha)^2 = r_3^2 \sin^2 \beta + r_3^2 \cos^2 \beta = r_3^2.$$

Transposing terms gives

$$2r_1 r_2 \cos \alpha = r_1^2 + r_2^2 - r_3^2.$$

A similar argument yields

$$2r_1 r_3 \cos \beta = r_1^2 + r_3^2 - r_2^2,$$

$$2r_2 r_3 \cos \gamma = r_2^2 + r_3^2 - r_1^2.$$

This extremum is a maximum, so we have shown that

$$2r_1 r_2 \cos \alpha + 2r_1 r_3 \cos \beta + 2r_2 r_3 \cos \gamma \leq r_1^2 + r_2^2 + r_3^2.$$

Now, it is known (see Bottema et al., *Geometric Inequalities*, Groningen, 1969, p. 23, no. 2.20) that if x, y, z are arbitrary positive numbers, then

$$x \cos \alpha + y \cos \beta + z \cos \gamma \leq \frac{yz}{2x} + \frac{zx}{2y} + \frac{xy}{2z}$$

with equality if and only if

$$\frac{1}{x} : \frac{1}{y} : \frac{1}{z} = a : b : c,$$

where a, b, c denote the sides of ABC . Therefore

$$\sum R_i^2 \sin^2 \alpha \leq 3 \sum r_i^2$$

with equality if and only if

$$(*) \quad r_1 : r_2 : r_3 = a : b : c.$$

But it is known that $(*)$ holds if and only if P is the symmedian point of ABC . (See, e.g., R. A. Johnson, *Modern Geometry*, Boston, 1929, p. 214.) This completes the proof.

Also solved by Leon Bankoff, M. G. Greening (Australia), Hans Kappus (Switzerland), M. S. Klamkin, F. Leuenberger (Switzerland), Simeon Reich (Israel), C. S. Venkataraman (India), and the proposer.

Non-Negative Forms

E 2348 [1972, 304]. *Proposed by Leonard Carlitz, Duke University*

Let P be a point in the interior of a triangle ABC . Let R_1, R_2, R_3 denote the distances from P to the vertices of ABC and let r_1, r_2, r_3 denote the perpendicular distances from P to the sides of ABC . Show that

$$(1) \quad \Sigma R_1(r_1 + r_3) \geq \Sigma(r_1 + r_2)(r_1 + r_3),$$

$$(2) \quad \Sigma(R_1 + R_2)(R_1 + R_3) \geq 4 \Sigma(r_1 + r_2)(r_1 + r_3),$$

with equality if and only if ABC is equilateral and P is its center.

Solution by M. S. Klamkin, Ford Scientific Laboratory. To satisfy (1) we prove a stronger inequality. For the triangle ABC let a, b, c be the lengths of the sides BC, CA, AB , respectively. From [1, p. 107] we have

$$(3) \quad R_1 \geq \frac{r_2c + r_3b}{a}, \quad R_2 \geq \frac{r_1c + r_3a}{b}, \quad R_3 \geq \frac{r_1b + r_2a}{c},$$

with equality if and only if ABC is equilateral and P is its center. We now prove that

$$(4) \quad \Sigma a^{-1}(r_2c + r_3b)(r_2 + r_3) \geq \Sigma(r_1 + r_2)(r_1 + r_3).$$

This inequality implies (1). This inequality is actually valid for all real r_1, r_2, r_3 since it will be shown to be a non-negative quadratic form with equality if and only if $a = b = c$. The matrix associated with (4) is given by

$$M = \begin{bmatrix} \frac{b^2 + c^2 - bc}{bc} & \frac{a + b - 3c}{2c} & \frac{a + c - 3b}{2b} \\ \frac{b + a - 3c}{2c} & \frac{c^2 + a^2 - ca}{ca} & \frac{b + c - 3a}{2a} \\ \frac{c + a - 3b}{2b} & \frac{c + b - 3a}{2a} & \frac{a^2 + b^2 - ab}{ab} \end{bmatrix}$$

As is well known, (4) is a non-negative form if the three principal minors M_1, M_2, M_3 of M are non-negative. After some algebraic manipulation, we find that

$$bc M_1 = (b - c)^2 + bc > 0,$$

$$4abc^2 M_2 = 4c^2(\Sigma a^2 - \Sigma ab) + ab(2 \Sigma ab - \Sigma a^2) > 0, \text{ and}$$

$$(x + y)^2(y + z)^2(z + x)^2 M_3 = (\Sigma xy)(\Sigma x^2 - \Sigma xy)(\Sigma x^2 + 3 \Sigma xy) \geq 0$$

with equality if and only if $x = y = z$, or equivalently $a = b = c$. Here we simplify the calculation of M_3 by using the duality transformation [2]

$$a = y + z, \quad b = z + x, \quad c = x + y,$$

where x, y, z are arbitrary non-negative numbers, not all zero.

Inequality (2) follows from adding the following two inequalities found in [1, p. 110]:

$$3 \sum R_2 R_3 \geq 12 \sum r_2 r_3, \quad \sum R_1^2 \geq 4 \sum r_1^2.$$

These inequalities are equalities if and only if ABC is equilateral and P is its center.

1. O. Bottema et al., *Geometric Inequalities*, Noordhoff, Groningen, 1969.

2. M. S. Klamkin, *Duality in triangular inequalities*, Ford Motor Company, preprint, July 1971.

Also solved by G. V. Ferrer (Mexico), F. Leuenberger (Switzerland), Michael Goldberg, and the proposer.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers — The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before June 30, 1973. Contributors (in the United States) who desire acknowledgement of receipt of their solutions are asked to enclose self-addressed, stamped postcards.

5883 [1972, 1042]. *Proposed by Frank Bernhart, Kansas State University*

Correction. The condition that set S be finite was inadvertently omitted from the hypothesis.

5900. *Proposed by E. R. Gentile, University of Buenos Aires, Argentina*

Let A and B be abelian groups (or modules over a principal ideal domain) such that $A \otimes B$ is a nonzero free abelian group (module). Prove that A and B are free.

5901. *Proposed by E. D. Dixon, Tennessee Technological University*

If a and x are elements of a ring R we denote $[a, x] = [a, x]_1 = ax - xa$ and, in general, $[a, [a, x]_h] = [a, x]_{h+1}$ for all positive integers h . Show that if P is a polynomial with coefficients which are integers or coefficients which are in the center of R , then

$$(1) [a, [P(a), x]_h] = [P(a), [a, x]_h] \text{ and}$$

$$(2) \text{ if } [a, x]_h = 0 \text{ then } [P(a), x]_h = 0.$$

5902. *Proposed by John H. Hubbard*

Integration is with respect to Lebesgue measure. Let $X \subset \mathbb{R}^m$ be a measurable set of finite measure. For any function $f: \mathbb{R}^m \rightarrow \mathbb{R}$, call $X_f^+ = \{x \in X: f(x) \geq 0\}$, $X_f^- = \{x \in X: f(x) \leq 0\}$.

Prove that for any whole number $n \geq 1$, there exists a polynomial function $p: \mathbb{R}^m \rightarrow \mathbb{R}$ of degree at most n such that for any polynomial function $q: \mathbb{R}^m \rightarrow \mathbb{R}$ of degree at most n satisfying $q(0) = 0$,

$$\int_{x_p^+} q = \int_{x_p^-} q.$$

5903. *Proposed by G. A. Heuer, Concordia College, and Albert Wilansky, Lehigh University*

If B is a two-dimensional noncommutative algebra over R (the real numbers), it is known that the multiplication in B is given by $ab = f(a)b$ or by $ab = f(b)a$ for some linear functional f on B . (Cf. Wilansky, *Functional Analysis*, problem 40, p. 258.) Is there a noncommutative multiplication in the set R^2 which, together with the usual vector addition makes R^2 into an associative ring which is not an algebra?

5904. *Proposed by S. Y. Chen and S. C. Hsieh, National Tsing Hua University, Taiwan*

Consider a semigroup S in which each pair of elements a, b satisfy $aba = a$. For such a in S let T_a be the set of elements b such that $ab = a$. Show that if, for any a , T_a contains more than half the elements of S then $T_b = S$ for all b in S .

5905. *Proposed by J. G. Wendel, University of Michigan*

Let x be a mapping of $(0, 1)$ into Euclidean space R^d and let $\{u_n\}$ be a countable dense set of vectors on the unit sphere of R^d . Suppose that for each n , $\limsup_{t \rightarrow 0} u_n \cdot x(t) = 1$. (a) Prove that $\limsup_{t \rightarrow 0} \|x(t)\| = 1$. (b) Can the denseness of the set $\{u_n\}$ be dispensed with? (c) Is the result true in Hilbert space?

SOLUTIONS OF ADVANCED PROBLEMS

Generalized Bases in Cross Spaces

5826 [1971, 1143]. *Proposed by Richard Tapia, Rice University*

Let X be a Banach space with dual X^* . Consider $B = \{(x_\alpha, y_\alpha): \alpha \in A\} \subset X \times X^*$ such that

$$(1) \quad \langle x_\alpha, y_\beta \rangle = \delta_{\alpha\beta} \text{ (Kronecker delta).}$$

Arsove and Edwards call B a generalized basis if in addition to (1) we have

$$(2) \quad \text{The linear span of } \{x_\alpha: \alpha \in A\} \text{ is dense in } X,$$

Davis calls B a dual generalized basis if in addition to (1) we have

- (3) The linear span of $\{y_\alpha: \alpha \in A\}$ is dense in X^* .

Prove or disprove: these two notions of bases coincide in Hilbert space.

Solution by E. M. Klein, University of Wisconsin — Milwaukee. We disprove the assertion. In the Hilbert space l^2 let $x_n = (0, 0, \dots, 0, 1, 0, \dots)$ where the 1 is in the $(n+1)$ -th position and let $y_n = (1, 0, \dots, 0, 1, 0, \dots)$ where the second 1 is in the $(n+1)$ -th position. Since l^2 is its own dual, the system $B = \{(x_n, y_n): n = 1, 2, 3, \dots\}$ satisfies (1). Equation (2) is not satisfied since the vector $(1, 0, 0, 0, \dots)$ is orthogonal to all the x_n . However, (3) is satisfied since if $b = (b_1, b_2, b_3, \dots)$ is orthogonal to all the y_n , then $b_n = -b_1$ for $n = 2, 3, \dots$ and b_n converges to 0, so $b = (0, 0, 0, \dots)$. Hence B is a dual generalized basis but not a generalized basis, and

$$B' = \{(y_n, x_n): n = 1, 2, 3, \dots\}$$

is a generalized basis but not a dual generalized basis.

Also solved by J. W. Evans, R. B. Israel, A. A. Jagers (Netherlands), and P. J. Owens (England).

Automorphisms of p -Groups

5827 [1971, 1144]. *Proposed by Ronald Hirshon, Polytechnic Institute of Brooklyn*

Let B be a finite p group, $p > 2$. Let Z , the center of B , be cyclic with w a generator of Z . Does there exist an automorphism ε of B such that $w\varepsilon = w^j$ for some j with $j \not\equiv 1 \pmod{p}$? If the answer is not always yes, will ε exist if we assume there is a maximal subgroup of B not containing Z ?

Editorial Note. No solution has been received for this problem. The proposer indicates that he has verified a "yes" answer to the first question for groups of order p^n with $n \leq 5$. In correspondence with the editors, J. L. Alperin says that the answer to the first question is "no" for $p \geq 7$, and that there should exist (possibly tedious) counterexamples for all $p > 2$. Alperin also conjectures that the answer is "no" for the second question. Further contributions are invited.

Representations by Permuting Countable Subsets of $(0,1)$

5828 [1971, 1144]. *Proposed by D. P. Giesy, Western Michigan University*

Let $x \in (0, 1)$. Is there an enumeration $\{q_n\}$ of the rationals in $(0, 1)$ such that $\sum_{n=1}^{\infty} q_n/2^n = x$? (See Problem 5700 [1970, 1018], especially the Editorial Note.)

5829 [1971, 1144]. *Proposed by D. P. Giesy, Western Michigan University*

Q is a countable subset of $(0, 1)$. Find necessary and sufficient conditions on Q that it have the property: For every $x \in (0, 1)$ there exists an enumeration q_1, q_2, \dots

of Q such that $\sum_{n=1}^{\infty} q_n/2^n = x$. (See Problem 5700 [1970, 1018], especially the Editorial Note, also the preceding problem.)

Solution by Neal Felsinger, Edgewood Arsenal, Maryland. It is obvious that $\inf Q = 0$ and $\sup Q = 1$ are necessary. We will show that they are also sufficient. Let $\{q_m\}$ be an ordering of Q and suppose $x \in (0, 1)$ is given. We will define sequences $\{x_k\}_{k=1}^{\infty}$, $\{y_n\}_{n=0}^{\infty}$ such that $y_n = \sum_{k=1}^n x_k/2^k$, $x_k \in Q$ and $x > y_n > x - 2^{-n}$. We set $y_0 = 0$. Suppose y_n is defined and let m be the least positive integer such that q_m is not an x_k , $k \leq n$. We consider three cases:

(i) $x > y_n + q_m/2^{n+1} > x - 2^{-(n+1)}$: Merely let $x_{n+1} = q_m$, $y_{n+1} = y_n + q_m/2^{n+1}$.

(ii) $y_n + q_m/2^{n+1} \geq x$: Let j be the unique integer such that $y_n + q_m/2^{j+1} < x$ and $y_n + q_m/2^j \geq x$. Let $x_{j+1} = q_m$ and let x_k , $n < k \leq j$, be any element of Q not already chosen, less than $x - (y_n + q_m/2^{j+1})$. Then

$$y_{j+1} = (y_n + q_m/2^{j+1}) + \sum_{k=n+1}^j x_k/2^k < (y_n + q_m/2^{j+1}) + (x - (y_n + q_m/2^{j+1})) = x$$

and

$$\begin{aligned} x_j - 2^{-(j+1)} &\leq y_n + q_m/2^j - 2^{-(j+1)} = (y_n + q_m/2^{j+1}) + (q_m - 1)/2^{j+1} \\ &< y_n + q_m/2^{j+1} < y_{j+1}. \end{aligned}$$

(iii) $y_n + q_m/2^{n+1} \leq x - 2^{-(n+1)}$: Let $j > n$ be the unique integer satisfying

$$y_n + \sum_{k=n+1}^j 2^{-k} \leq x \text{ and } y_n + \sum_{k=n+1}^{j+1} 2^{-k} > x.$$

For $n < s \leq j$, determine x_s in Q , not already selected, so that

$$x_s > 1 - \left[y_n + \sum_{k=n+1}^j 2^{-k} - (x - 2^{-(j+1)}) \right],$$

so that $x_s \geq 1 - 2^{-(j+1)}$. Then

$$x - 2^{-j} < x - 2^{-(j+1)} < y_n + \sum_{k=n+1}^j x_k/2^k < y_n + \sum_{k=n+1}^j 2^{-k} \leq x.$$

Thus $x - 2^{-j} < y_j < x$ and $x - 2^{-(j+1)} < y_j < y_j + q_m/2^{j+1}$. Therefore either case (i) or case (ii) applies to y_j depending on whether or not $y_j + q_m/2^{j+1} < x$.

Finally $x = \lim y_n = \sum_{k=1}^{\infty} x_k/2^k$ and $Q = \{x_k\}_{k=1}^{\infty}$.

Also solved by R. L. Enison, P. J. Owens (England), and the proposer. Problem 5828 only solved by G. J. Butler and by R. A. Struble.

Editorial Note. An earlier solution to Problem 5828 was part of the solution to Problem 5700, given by the proposers of 5700, R. A. Struble and R. E. Chandler. Actually Struble had also proposed 5828 independently.

Function with a Natural Boundary

5830 [1971, 1144]. *Proposed by Leonard Carlitz and R. A. Scoville, Duke University*

Let α be a positive irrational number and put

$$\phi(z) = \sum_{n=1}^{\infty} z^{[\alpha n]},$$

where $[\alpha n]$ denotes the greatest integer $\leq \alpha n$. Show that $\phi(z)$ has the unit circle for a natural boundary.

Solution by C. C. Rousseau, Memphis State University. A theorem due to Szegő (Dienes, *The Taylor Series*, p. 324) states that if

$$f(z) = \sum_{k=0}^{\infty} a_k z^k$$

and if $\{a_k\}$ contains only finitely many distinct numbers, then $f(z)$ has the unit circle as a natural boundary unless $a_{k+p} = a_k$ for all sufficiently large k , in which case $f(z)$ is the rational function, $P(z)/(1 - z^p)$.

Writing $\phi(z)$ as a power series

$$\phi(z) = \sum_{k=0}^{\infty} a_k z^k,$$

where a_k is the number of integers $n \geq 1$ such that $[\alpha n] = k$, we see that, for fixed α , $\{a_k\}$ contains at most finitely many distinct numbers. Thus, we need only show that if α is irrational, $\{a_k\}$ cannot be periodic.

Assume that $\{a_k\}$ is periodic, i.e., that there exists p such that $a_{k+p} = a_k$ for all sufficiently large k . Using the assumption of periodicity and setting

$$q = \sum_{j=0}^{p-1} a_{k+j}$$

it follows that if n_k is the smallest integer n such that $[\alpha n] = k$, then $n_k + mq$ is the smallest integer n such that $[\alpha n] = k + mp$. Hence, we write

$$\alpha n_k = k + \beta_k \text{ and } \alpha(n_k + mq) = k + mp + \beta_{k+mp},$$

where $0 \leq \beta_{k+mp} < \alpha$ ($m = 0, 1, \dots$). From the difference of these two expressions we obtain

$$\beta_{k+mp} = \beta_k + mq(\alpha - p/q).$$

It follows that $0 \leq \beta_{k+mp} < \alpha$ cannot be satisfied for all m unless $\alpha = p/q$. Hence, if α is irrational, $\{a_k\}$ cannot be periodic.

Also solved by L. W. Carroll, L. Kuipers, and the proposers.

Carroll's proof is immediate from a theorem of Hecke on the fact that, for β irrational, $\Sigma[(k+1)\beta]z^k$ has a natural boundary. Kuiper's proof is based on Problem 168 in Vol. 1 of Polya-Szegő's *Aufgaben*, a result equivalent to Hecke's theorem.

Convex Sets with Nonempty Interiors

5831 [1971, 1144]. *Proposed by Albert Wilansky, Lehigh University*

Let C be a convex closed set in a normed space such that $C + D_1 \supset D_{1+\varepsilon}$. Must C have nonempty interior? (Here $D_r = \{x: \|x\| \leq r\}$, $\varepsilon > 0$.)

Solution by R. B. Israel, University of Chicago. C must have nonempty interior and, in fact, we must have $D_\varepsilon \subset C$. For real numbers r we define $rC = \{rc: c \in C\}$. By the convexity of C it is easy to see that for positive r and s we have $rC + sC = (r+s)C$. Now we have $D_{1+\varepsilon} \subset D_1 + C$, so

$$D_{(1+\varepsilon)^2} = (1+\varepsilon)D_{1+\varepsilon} \subset D_{1+\varepsilon} + (1+\varepsilon)C \subset D_1 + (1+(1+\varepsilon))C.$$

By iterating n times, we find that

$$\begin{aligned} D_{(1+\varepsilon)^n} &\subset D_{(1+\varepsilon)^{n-1}} + (1+\varepsilon)^{n-1}C \subset \dots \\ &\subset D_1 + \sum_{j=0}^{n-1} (1+\varepsilon)^j C = D_1 + \varepsilon^{-1}((1+\varepsilon)^n - 1)C, \end{aligned}$$

and, letting $(1+\varepsilon)^n = r$, we have $D_\varepsilon \subset D_{\varepsilon/r} + (1-1/r)C$. Let x be any member of D_ε . Then for each n there are y_n and z_n such that $x = y_n + z_n$ with $y_n \in D_{\varepsilon/r}$ and $r(r-1)^{-1}z_n \in C$. But as $n \rightarrow \infty$, $r \rightarrow \infty$, so $y_n \rightarrow 0$ and $z_n \rightarrow x$. Moreover, $r(r-1)^{-1}z_n \rightarrow x$, and since C is closed we must have $x \in C$. Thus $D_\varepsilon \subset C$.

Also solved by P. R. Chernoff, Moshe Feder & Simeon Reich (Israel), John Horvath, R. M. Koch, L. E. Mattics, P. J. Owens (England), and the proposer.

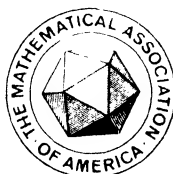
Notes. (1) The proposer states that the problem arises in providing an elementary proof that a map between Banach spaces is almost open. This question arises in determining a sufficient condition for such a map to be onto. See Bade and Curtiss, *Pacific J. of Math.* 18(1966), pp. 391, ff, Theorem 1.2; also Kaufman, *Proceedings of the A. M. S.*, 17 (1966), pp. 767-768.

(2). Koch rephrases the problem and proves: *Let C be a convex set in a topological vector space such that $C + U \supset (1+\varepsilon)U$, where U is a bounded neighborhood of 0, $\varepsilon > 0$. Then the closure of C is a neighborhood of C , actually $\bar{C} \supset \varepsilon U$.*

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF
THE MATHEMATICAL ASSOCIATION OF AMERICA



VOLUME 80

NUMBER 4

CODEN: AMMYAE

CONTENTS

A Unified Theory of Integration	E. J. McSHANE	349
Highlights in the History of Spectral Theory	L. A. STEEN	359
The Legend of John von Neumann	P. R. HALMOS	382
Alternating Euler Paths for Packings and Covers	C. T. ZAHN, JR.	395

MATHEMATICAL NOTES

Stable Laws and the Imbedding of L^p Spaces	MAREK KANTER	403
A Convex Matrix Function	M. H. MOORE	408
Solution of Fejes Tóth's Illumination Problem	B. R. HENRY	409
A Covering Theorem	J. C. KIEFFER	410
Distributivity over the Dirichlet Product and Completely Multiplicative Arithmetical Functions	ERIC LANGFORD	411
Perfect Parallelograms	R. W. SIELAFF	414
A Crowded Set of Non-intersecting Lines	J. A. EIDSWICK	415

RESEARCH PROBLEMS

A Deception Game	JOEL SPENCER	416
----------------------------	--------------	-----

CLASSROOM NOTES

Traffic Flow: Laplace Transforms	E. A. BENDER AND L. P. NEUWIRTH	417
Irrational Numbers	J. P. JONES AND S. TOPOROWSKI	423
A Simple Proof of the Formula $\sum_{k=1}^{\infty} k^{-2} = \pi^2/6$	IOANNIS PAPADIMITRIOU	424
Another Elementary Proof of Euler's Formula for $\zeta(2n)$	T. M. APOSTOL	425

(Continued on inside cover)

APRIL

1973

MATHEMATICAL EDUCATION

An Integrated Sequence in the Mathematical Sciences for Undergraduate	
Business Students	R. H. RANGLES AND A. J. SCHAEFFER 431
ELEMENTARY PROBLEMS AND SOLUTIONS	433
ADVANCED PROBLEMS AND SOLUTIONS	440
REVIEWS	447
NEWS AND NOTICES	466
MATHEMATICAL ASSOCIATION OF AMERICA	467
October Meeting of the North Central Section	467
Officers and Committees as of February 1, 1973.	468
Calendars of Future Meetings	474

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss.

Backlog: Main Articles 12 months, Math. Notes 15 months, Research Problems 7 months, Classroom Notes 11 months, Math. Education 10 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to ALEX ROSENBERG, Department of Mathematics, Cornell University, Ithaca, N.Y. 14850; NOTES, etc.: to the corresponding Associate Editor: ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D.C. 20036.

HARLEY FLANDERS, *Editor*

ALEX ROSENBERG, *Editor-Elect*

ASSOCIATE EDITORS

JOSHUA BARLAZ
E. R. BERLEKAMP
JANE W. DI PAOLA
ROBERT GILMER
RICHARD GUY
RAOUL HAILPERN

J. G. HARVEY
ERIC S. LANGFORD
P. D. LAX
ARTHUR MATTUCK
M. W. POWNALL
GIAN-CARLO ROTA

SEYMOUR SCHUSTER
J. ARTHUR SEEBACH, Jr.
E. P. STARKE
LYNN A. STEEN
JAMES WENDEL

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are \$12.50. For nonmembers the subscription price is \$18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices.

Copyright © The Mathematical Association of America (Incorporated), 1973

PRINTED IN THE UNITED STATES OF AMERICA

A UNIFIED THEORY OF INTEGRATION

E. J. McSHANE, University of Virginia

1. Introduction. An undergraduate student of mathematics, or science, or engineering, is usually introduced to an assortment of integrals. First he meets the one-dimensional Darboux integral, defined in terms of the elementary integrals of step-functions above and below the integrand. By the time he finishes a more advanced course in calculus he has met multiple integrals, several kinds of “improper” integrals, line integrals, etc. But, beyond this, in mathematics or theoretical physics or engineering science, it is important to know (or at least be willing to believe in!) the Lebesgue integral. Unfortunately, few students of science or engineering can afford the time to start afresh with a new theory of integration, especially since it appears abstract and disconnected from all that they have previously learned. Even students of mathematics not specializing in analysis often fail to develop any ease in using the Lebesgue integral. (More than once I have found graduate students unable to answer the question “what is the Lebesgue integral from 0 to 1 of $x^2 dx$?”)

Since this lack of unity in integration theory is an inconvenience to all sorts of students and hinders the access of scientists to parts of mathematics that are important in theoretical physics and engineering science, I am enthusiastic about promoting the use of a theory of integration in which there is in effect only one definition of the integral. This is presented in successively more general settings, but each change is merely a rather simple amendment to take in more territory. We never have to abandon it in favor of the Lebesgue integral, because it is equivalent to the Lebesgue integral. For use in teaching, the crucial point is to show that this theory can be smoothly fitted on to existing classroom methods at any point, but preferably just after the student has finished a beginning course in which the Darboux integral has been defined, and its importance has been demonstrated by examples, and the student has acquired some facility in its use by exercises. The following pages, therefore, will be devoted to showing how the integral can be presented to such a beginner and can be extended step by step to satisfy all the requirements of undergraduate

E. J. McShane received his Ph.D. from the University of Chicago in 1930. His dissertation, on existence theorems for isoperimetric problems in the calculus of variations, was written under the direction of Gilbert Ames Bliss. He has been a National Research Council Fellow (1930–32); a Hilfsassistent at Göttingen; on the Princeton faculty, 1933–35; and since then a professor at the University of Virginia, with some interruptions — the war years at the Ballistic Research Laboratory in the Aberdeen Proving Ground, a year at the Institute for Advanced Study, a year at the University of Utrecht, a year at the Rockefeller University, a semester at the University of Kyoto. He is a member of the National Academy of Sciences and of the American Philosophical Society, and has been president of the MAA and of the AMS. The Association has awarded him the Chauvenet Prize and the Award for Distinguished Service. He has published numerous papers, mostly on calculus of variations, integration theory, control theory and stochastic processes; also several books, including one entitled *Integration* which this paper is intended to sabotage. (*Submitted by author.*)

mathematics, without asking too much of reasonably capable undergraduates. It has already been shown (McShane [1]) that the same type of definition can be used, and deep theorems proved, in settings of great generality. Our concern here is with the quite different, and in my opinion much more important, problem of making modern integration theory painlessly available to a large body of students.

We therefore suppose that we are facing a group of students who know the definition of the integral of a step-function on an interval in one-dimensional space R^1 , and who know that the (Darboux) integral

$$\int_a^b f(x)dx$$

is the unique number J which is the infimum of the integrals of step-functions $S \geq f$ and the supremum of the integrals of step-functions $s \leq f$, provided that there is such a unique J . In this whole presentation we shall be forced to omit a great deal and condense the rest much more than would be appropriate in a classroom, so that we ask much of the reader in filling in details and even believing unsupported assertions. The full presentation will be in a book whose manuscript is still less than half finished.

2. Transition from the Darboux integral. In one-dimensional space R^1 a **neighborhood** of a point \bar{x} will mean an open interval that contains \bar{x} . We shall often make use of functions δ defined on some set E in R^1 (or, later, in some other space) such that for each \bar{x} in E , $\delta(\bar{x})$ is a neighborhood of \bar{x} . Such a neighborhood-valued function will be called a **gauge** on E . For example, we can phrase the definition of continuity thus. A real-valued function f on E is continuous on E if for each positive ε there is a gauge δ such that whenever $\bar{x} \in E$ and $x \in E \cap \delta(\bar{x})$, it is true that $|f(x) - f(\bar{x})| < \varepsilon$.

We introduce two more expressions. A **partition** of the right-closed interval $(a, b]$ will be a finite set $P = \{(\bar{x}_1, A_1), \dots, (\bar{x}_k, A_k)\}$ of pairs in which each \bar{x}_i is a point of the closed interval $[a, b]$, each A_i is a right-closed subinterval of $(a, b]$, and each point of $(a, b]$ belongs to exactly one of the A_i . The partition P , or more generally, any collection of pairs (\bar{x}_i, A_i) , is said to be **δ -fine** if for each i , the interval A_i is contained in the neighborhood $\delta(\bar{x}_i)$ of the point \bar{x}_i .

Suppose now that f is defined and continuous on a closed interval $[a, b]$. Let ε be positive. There is a gauge δ on $[a, b]$ such that if $\bar{x} \in [a, b]$ and $x \in [a, b] \cap \delta(\bar{x})$ then $|f(x) - f(\bar{x})| < \varepsilon/2(b - a)$. Let $P = \{(\bar{x}_1, A_1), \dots, (\bar{x}_k, A_k)\}$ be any δ -fine partition of $(a, b]$; for the moment we merely assume that such partitions exist. If x is any point in the interval A_i , x is in $\delta(\bar{x}_i)$ because $A_i \subset \delta(\bar{x}_i)$. So, by the choice of gauge δ , we have

$$(2.1) \quad f(\bar{x}_i) - \varepsilon/2(b - a) < f(x) < f(\bar{x}_i) + \varepsilon/2(b - a).$$

We now define two step-functions s, S on $(a, b]$ by setting

$$s(x) = f(\bar{x}_i) - \varepsilon/2(b-a), \quad S(x) = f(\bar{x}_i) + \varepsilon/2(b-a)$$

for all x in A_i ($i = 1, \dots, k$). Then by (2.1) we have $s < f < S$. If we use the symbol $m A_i$ to denote the length of A_i , by elementary calculus we have

$$\begin{aligned} (2.2) \quad \int_a^b S(x)dx &= \sum_{i=1}^k \{f(\bar{x}_i) + \varepsilon/2(b-a)\} m A_i \\ &= \sum_{i=1}^k f(\bar{x}_i) m A_i + [\varepsilon/2(b-a)] \sum m A_i \\ &= \sum_{i=1}^k f(\bar{x}_i) m A_i + \varepsilon/2, \end{aligned}$$

and similarly

$$(2.3) \quad \int_a^b s(x)dx = \sum_{i=1}^k f(\bar{x}_i) m A_i - \varepsilon/2.$$

Therefore there are step-functions s, S such that $s < f < S$ and the integrals of s and S differ by the arbitrarily small number ε . By elementary calculus, f has an integral; we denote it by J . Moreover, J is between the integrals of s and S . But by (2.2) and (2.3) this implies that

$$\sum_{i=1}^k f(\bar{x}_i) m A_i - \varepsilon/2 \leq J \leq \sum_{i=1}^k f(\bar{x}_i) m A_i + \varepsilon/2,$$

whence

$$(2.4) \quad \left| \sum_{i=1}^k f(\bar{x}_i) m A_i - J \right| < \varepsilon.$$

We now have exhibited a new procedure for identifying the Darboux integral J of the continuous function f . It is the (unique) number J such that to each positive ε there corresponds *at least one* gauge δ such that *for every* δ -fine partition $P = \{(\bar{x}_1, A_1), \dots, (\bar{x}_k, A_k)\}$ of $(a, b]$, inequality (2.4) is satisfied. Since this procedure always gives back the known Darboux integral for continuous f , we have the privilege of using it as the definition of the integral for such f . But the same procedure does much more. It picks out one number J (by inequality (2.4)) for many discontinuous and even unbounded f , for which no Darboux integral exists. When this happens, we still apply the name "integral" to the number J that the procedure selects for us, thus.

DEFINITION 2.1. *A real-valued function f is said to be integrable over a right-closed interval $(a, b]$ if it is defined on the closure $[a, b]$, and there is a number J such that: to each positive ε there corresponds a gauge δ such that for every δ -fine partition $P = \{(\bar{x}_1, A_1), \dots, (\bar{x}_k, A_k)\}$ of $(a, b]$ it is true that*

$$\left| \sum_{i=1}^k f(\bar{x}_i) m A_i - J \right| < \varepsilon.$$

In this case we define

$$\int_a^b f(x) dx = J.$$

In the theory of the Darboux (or Riemann) integral, to prove the integrability of a function on $[a, b]$ we have to show the existence of step-functions s and S like those of the preceding paragraph. This can be accomplished by proving and using the Heine-Borel theorem. In order to develop the theory of the integral defined in Definition 2.1, we need to know that for every gauge δ on $[a, b]$ there exist δ -fine partitions of $[a, b]$ —even though most undergraduates will falsely regard this as evident. The proof is almost identical with that of the Heine-Borel theorem.

It is easy to prove that step-functions are integrable, their integrals being given by the familiar formula. From this we deduce that every Darboux-integrable function ($f(x): a \leq x \leq b$) is also integrable by Definition 2.1, the integrals being the same. So we have not abandoned the Darboux integral; we have generalized it. But the generalization is considerable. For example, if $f(x) = x^{-\frac{1}{2}}$ for $x > 0$ and $f(x) = 0$ for $x \leq 0$, by exhibiting a gauge δ for each positive ε we can show that

$$(2.5) \quad \int_0^b f(x) dx = 2b^{\frac{1}{2}} \quad (b > 0).$$

This f is not Darboux integrable, since there is no step-function above f . Equation (2.5) is more easily established after we prove the monotone convergence lemma (§3), which will in fact enable us to prove that all the “absolutely convergent improper integrals” of advanced calculus are covered by Definition 2.1, with no “impropriety” at all.

Definition 2.1 specifies the integral as a limit, in a known generalized formulation of limit-theory (see, for example, [2], pp. 7–29). So uniqueness, the Cauchy criterion, etc., all hold. But it is probably better for normal advanced-calculus students just to prove these statements without wandering off into abstract limit theory.

Some elementary texts use Riemann’s definition instead of that of Darboux. The last equation in Definition 2.1 is defined to mean that to each positive ε there corresponds a positive constant δ such that for every partition $\{(\bar{x}_1, A_1), \dots, (\bar{x}_k, A_k)\}$ with each \bar{x}_i in A_i and each A_i with length less than δ , it is true that

$$\left| \sum_{i=1}^k f(\bar{x}_i) m A_i - J \right| < \varepsilon.$$

A student familiar with this definition should be able to make the transition to Definition 2.1 with ease. But any student mature enough to master Riemann’s def-

inition should be able to comprehend Definition 2.1, and he would profit by replacing Riemann's definition by Definition 2.1 in the first place.

3. Absolute integrability and monotone convergence. The integral defined in (5) has the same elementary properties as the Riemann integral: it is linear in f ; it is non-negative when f is non-negative; f is integrable over $(c, d]$ when it is integrable over an interval $(a, b]$ that contains $(c, d]$; it is integrable over $(a, b]$ when it is integrable over each of two disjoint intervals $(a, c], (c, b]$ whose union is $(a, b]$, and then

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx;$$

and the familiar differentiability properties hold.

Because Definition 2.1 closely resembles the definition of the Riemann integral, the customary proofs can be carried over with little change, and there is no point in dwelling on them. However, we shall give condensed proofs of two statements absent from the elementary theory.

Corresponding to a partition $P = \{(\bar{x}_1, A_1), \dots, (\bar{x}_k, A_k)\}$ of $(a, b]$ and a function f on $[a, b]$ the sum

$$\sum_{i=1}^k f(\bar{x}_i) m A_i$$

occurs frequently. It is convenient to call it a "Riemann sum" and to assign it a symbol $S(P; f)$.

We now prove that if f is integrable over $(a, b]$ it has a property that I have named "**absolute integrability**".

THEOREM 3.1. *Let f be integrable over $(a, b]$ and let J denote its integral. Let ε be positive, and let δ be a gauge on $[a, b]$ such that for every δ -fine partition P of $(a, b]$, $|S(P; f) - J| < \varepsilon$. Then for every two δ -fine partitions*

$$P' = \{(x'_1, A'_1), \dots, (x'_h, A'_h)\},$$

$$P'' = \{(x''_1, A''_1), \dots, (x''_k, A''_k)\}$$

of $(a, b]$, it is true that

$$\sum_{i=1}^h \sum_{j=1}^k |f(x'_i) - f(x''_j)| m(A'_i \cap A''_j) < 2\varepsilon.$$

Every point of $(a, b]$ belongs to exactly one of the intersections $A'_i \cap A''_j$, where (as throughout this proof) i ranges over $\{1, \dots, h\}$ and j over $\{1, \dots, k\}$. Define $x'_{ij} = x'_i$ and $x''_{ij} = x''_j$ if $f(x'_i) \geq f(x''_j)$; otherwise define $x'_{ij} = x''_j$ and $x''_{ij} = x'_i$. In either case,

$$(3.1) \quad f(x'_{ij}) - f(x''_{ij}) = |f(x'_i) - f(x''_j)|.$$

Both sets of pairs

$$P''' = \{(x'_{ij}, A'_i \cap A''_j)\}, \quad P^{iv} = \{(x''_{ij}, A'_i \cap A''_j)\}$$

are δ -fine partitions of $(a, b]$. So $S(P'''; f)$ and $S(P^{iv}; f)$ differ from J by less than ε , and therefore differ from each other by less than 2ε . That is,

$$\left| \sum_{i=1}^h \sum_{j=1}^k \{f(x'_{ij}) - f(x''_{ij})\} m(A'_i \cap A''_j) \right| < 2\varepsilon.$$

By (3.1) this is the conclusion of Theorem 3.1.

From this and the Cauchy criterion for existence of the integral, we can prove that if f is integrable over $(a, b]$, and L is a function defined on the set of values of f and satisfying $|L(y_1) - L(y_2)| \leq K|y_1 - y_2|$ (K a constant) for all y_1 and y_2 that are values of f , then $(L(f(x)): a \leq x \leq b)$ is integrable over $(a, b]$. As corollaries, if f, f_1 and f_2 are integrable over $(a, b]$, and f_1 and f_2 are bounded, the following functions are integrable: $f^+ = \max(f, 0)$, $f^- = \max(-f, 0)$, $|f|$, f_1^n ($n = 1, 2, 3, \dots$), $|f_1|^\alpha$ ($\alpha > 0$), $f_1 f_2$, and $1/f$ if f is bounded away from 0.

All the theorems mentioned so far have been valid for the Riemann (or Darboux) integral also. They apply to a larger class of functions; but if this were the only gain, it would not be enough to justify the introduction of Definition 2.1 into a calculus or advanced-calculus course. But now we shall prove a fundamentally important statement that has no analogue in the theory of the Darboux integral, namely a slightly restricted form of the monotone convergence theorem. In the proof we shall need the following rather easy consequence of Theorem 3.1.

THEOREM 3.2. *Let f, δ and ε be as in Theorem 3.1. Let $\{(\bar{x}_1, A_1), \dots, (\bar{x}_k, A_k)\}$ be a δ -fine set of pairs such that the A_i are pairwise disjoint right-closed subintervals of $(a, b]$ (the union of the A_i need not be all of $(a, b]$). Then*

$$\sum_{i=1}^k |f(\bar{x}_i) m A_i - \int_{A_i} f(x) dx| \leq 2\varepsilon.$$

THEOREM 3.3 (Monotone convergence lemma). *Let f_1, f_2, f_3, \dots be a sequence of real-valued functions all defined on $[a, b]$ and integrable over $(a, b]$. Assume that for each x in $[a, b]$, $f_1(x) \leq f_2(x) \leq f_3(x) \leq \dots$, and that $\lim_{n \rightarrow \infty} f_n(x)$ (which necessarily exists, finite or infinite) is actually finite. Then the limit-function, whose value at x is $\lim_n f_n(x)$, is integrable over $(a, b]$ if and only if the limit*

$$J = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx$$

is finite; and in that case,

$$\int_a^b [\lim_{n \rightarrow \infty} f_n(x)] dx = J.$$

Let g denote $\lim_n f_n$. There is no loss of generality in assuming $f_1 \geq 0$. Suppose $J < \infty$ (the "only if" part is trivial), and let ε be positive. We must exhibit a gauge δ on $[a, b]$ such that for every δ -fine partition P of $(a, b]$,

$$(3.2) \quad J - \varepsilon < S(P; g) < J + \varepsilon.$$

Define

$$(3.3) \quad J_n = \int_a^b f_n(x) dx.$$

For each n there is a gauge δ_n such that if P' is a δ_n -fine partition of $(a, b]$ then

$$(3.4) \quad |S(P'; f_n) - J_n| < \varepsilon/2^{n+2};$$

in addition, we may suppose $\delta_1(x) \supset \delta_2(x) \supset \delta_3(x) \supset \dots$ for all x in $[a, b]$. Define E_n to be the set of all x in $[a, b]$ such that

$$(3.5) \quad f_n(x) \geq [(2J + \varepsilon)/(2J + 2\varepsilon)]g(x).$$

Then $E_1 \subset E_2 \subset E_3 \subset \dots$. If $g(x) = 0$, $x \in E_1$, because then $f_1(x) = 0$. If $g(x) > 0$ the right member of (3.5) is less than the limit $g(x)$ of $f_n(x)$, so (3.5) holds for all large n . Hence $\bigcup E_n = [a, b]$.

The numbers J_1, J_2, \dots are non-decreasing and tend to J , so we can and do choose an N such that $J \geq J_N > J - \varepsilon/2$. Every x in $[a, b]$ belongs to exactly one of the sets

$$D_N = E_N, D_{N+1} = E_{N+1} \setminus E_N, \dots, D_m = E_m \setminus E_{m-1}, \dots.$$

If $\bar{x} \in D_n$, we define $\delta(\bar{x}) = \delta_n(\bar{x})$. Then δ is a gauge on $[a, b]$. Let

$$P = \{(\bar{x}_1, A_1), \dots, (\bar{x}_k, A_k)\}$$

be any δ -fine partition of $[a, b]$. Then P is δ_N -fine, so

$$\begin{aligned} S(P; g) &\geq S(P; f_N) \\ &> J_N - \varepsilon/2^{N+2}, \end{aligned}$$

and the first inequality in (3.2) holds.

For the other, let $I(n)$ be the set of those i in $\{1, \dots, k\}$ for which $\bar{x}_i \in D_n$ ($n = N, N+1, \dots$). There is a largest n , say n^* , for which $I(n)$ is not empty. By (3.5) and Theorem 3.2,

$$\begin{aligned} (3.6) \quad S(P; g) &= \sum_{n=N}^{n^*} \sum_{i \in I(n)} g(\bar{x}_i) m A_i \\ &\leq \sum_{n=N}^{n^*} \left\{ \sum_{i \in I(n)} [(2J + 2\varepsilon)/(2J + \varepsilon)] f_n(\bar{x}_i) m A_i \right\} \end{aligned}$$

$$\begin{aligned} &\leq [(2J + 2\varepsilon)/(2J + \varepsilon)] \sum_{n=N}^{n^*} \left\{ \sum_{i \in I(n)} \int_{A^i} f_n(x) dx + 2\varepsilon/2^{n+2} \right\} \\ &\leq [(2J + 2\varepsilon)/(2J + \varepsilon)] \left\{ \sum_{n=N}^{n^*} \sum_{i \in I(n)} \int_{A^i} f_{n^*}(x) dx + \sum_{n=N}^{n^*} \varepsilon/2^{n+1} \right\}. \end{aligned}$$

The union of the A_i , as i ranges over $I(n)$ and n ranges over N, \dots, n^* , is $(a, b]$. Also,

$$\sum_{n=N}^{n^*} \varepsilon/2^{n+1} < \sum_{n=1}^{\infty} \varepsilon/2^{n+1} = \varepsilon/2.$$

Therefore inequality (3.6) yields

$$\begin{aligned} S(P; g) &< [(2J + 2\varepsilon)/(2J + \varepsilon)] \left\{ \int_a^b f_{n^*}(x) dx + \varepsilon/2 \right\} \\ &\leq [(2J + 2\varepsilon)/(2J + \varepsilon)] \{J + \varepsilon/2\} \\ &= J + \varepsilon. \end{aligned}$$

So the second inequality in (3.2) holds, and the proof is complete.

The crucial idea of this proof is due to R. Henstock, who proved the analogous theorem for the "Riemann-complete" integral.

If f is non-negative and continuous on $[a, b]$ except at finitely many points x_1^*, \dots, x_h^* , we can define $f_n(x)$ to be 0 if x has distance less than $1/n$ from some x_j^* and to be $f(x)$ otherwise. From Theorem 3.3 we obtain the usual theory of the "improper" integral of f , which for us is merely the integral of f . By considering f^+ and f^- we obtain the theory of "absolutely convergent improper integrals".

4. Multidimensional integrals. Next, let a right-closed interval $(a, b]$ in R^n be defined to be the cartesian product of n right-closed intervals $(a^{(i)}, b^{(i)}]$ in R^1 , and let open intervals (a, b) and closed intervals $[a, b]$ be analogously defined. With this notation for $(a, b]$ we define $m(a, b] = (b^{(1)} - a^{(1)}) \dots (b^{(n)} - a^{(n)})$. Thus $m(a, b]$ is length if $n = 1$, area if $n = 2$, and volume if $n = 3$. If f is real-valued on $[a, b] \subset R^n$, the definition of

$$\int_{(a, b]} f(x) dx$$

is the same as Definition 2.1. All theorems generalize at once, except those involving differentiation. A new concept appears, namely integration by iteration. For this we can easily prove a form of Fubini's theorem adequate for the needs of advanced calculus. For simplicity we consider $n = 2$ and write (u, v) as another symbol for $x = (x^{(1)}, x^{(2)})$. If $A = (a, b] \times (c, d]$ and f is a step-function on $A^- = [a, b] \times [c, d]$, it is trivial that

$$(4.1) \quad \int_A f(x)dx = \int_c^d \left[\int_a^b f(u,v)du \right] dv.$$

Suppose that f is bounded on A^- , and that there is a closed subset F of A^- such that f is continuous on $A^- \setminus F$ and f as restricted to F is continuous on F . We first make the additional hypothesis that $f(x) \leq 0$ on F and $f(x) \geq 0$ on $A^- \setminus F$. For $j = 1, 2, 3, \dots$ we subdivide A^- into 4^j subintervals by dividing each side of A^- into 2^j intervals of equal length, and we define the step-function s_j by giving it on each subinterval the infimum of values of f on that subinterval. Then $s_1 \leq s_2 \leq \dots$, and (4.1) holds for each s_j . By several applications of the monotone convergence lemma we find that (4.1) holds for f . The additional hypothesis is easily removed. Moreover, the u and v need not be points of R^1 . The theorem holds for x, u, v in R^n, R^p, R^q respectively, where $p + q = n$.

If B is any set in any space, we define $\mathbf{1}_B$ to be the **characteristic function**, or **indicator function**, of B ; its value is 1 at all points of B and 0 at all other points of the space. If $\mathbf{1}_B$ is integrable over R^n , we say that B has finite measure, and we define

$$(4.2) \quad mB = \int_{R^n} \mathbf{1}_B(x)dx.$$

Technically, we have everything needed for a theory of measure. But the students lack the necessary mathematical maturity, and it is better to postpone measure theory to a later stage. However, (4.2) expresses the classical formula for "areas by double integration" and "volumes by triple integration," and (4.1) allows us to solve all the text-book problems by iterated integration.

5. Integration over unbounded sets. For the next extension we form \bar{R}^1 by adjoining $+\infty$ and $-\infty$ to the real system R^1 . These have the expected algebraic and order properties; we take

$$0 \cdot \infty = 0 \cdot (-\infty) = 0.$$

The neighborhoods of ∞ are the half-lines $\{x \in \bar{R}^1: x > a\}$ for all a in R^1 ; the neighborhoods of $-\infty$ are the half-lines $\{x \in \bar{R}^1: x < a\}$. Also, we include the half-lines $\{x \in R^1: x \leq b\}$ and $\{x \in R^1: x > a\}$ and R^1 itself among the right-closed intervals in R^1 , and for each of these unbounded intervals A we define $mA = \infty$. If f is defined on some subset of R^1 , $f(\infty)$ and $f(-\infty)$ are undefined. We give them the value 0. Now we can define integrals over unbounded intervals by Definition 2.1 without change. All theorems extend with at most minor changes; and the monotone convergence lemma provides us with the theory of "absolutely convergent improper integrals" over intervals (a, ∞) , $(-\infty, b)$ and $(-\infty, \infty)$. Again these integrals are not at all "improper". The limit process which in the customary calculus text defines the integral is for us merely a convenient device for computing an integral already defined and studied.

For brevity we omit discussion of three types of integral usually discussed in advanced calculus texts. These are integrals of vector-valued functions, to which Definition 2.1 applies if we understand the symbol $|\cdot|$ to denote the length of a vector; line integrals; and surface integrals, about which we have nothing significant to say.

At this stage the use of our integral in place of the Darboux integral has caused very little change in factual information or in the mental effort required to make proofs. But the student who has been familiarizing himself with the integral of Definition 2.1 now has the background material for further advances. Also, he will not be subjected to the trauma of being told to discard the integral with which he is familiar and change to the Lebesgue integral, for the integral of Definition 2.1 is the Lebesgue integral.

6. Probability measures. The use of the integral of §2 is particularly helpful to students in advanced undergraduate probability courses. Such courses demand more than the Darboux integral can provide. In many texts, discrete distributions are well treated; but the general case requires countably additive measures on σ -algebras of sets, and authors are caught between the dangers of overloading the students with measure theory and overloading the book with unproved (and sometimes untrue) assertions. Our integral removes this trouble.

We begin with an elementary probability measure P defined and non-negative on all right-closed intervals in R^n , such that $P(R^n) = 1$, and if a right-closed interval A is the union of pairwise disjoint right-closed intervals A_1, \dots, A_n then $P(A) = P(A_1) + \dots + P(A_n)$. We also add the **regularity** requirement that for each right-closed interval A and each positive ε , there is a right-closed interval B whose *closure* is contained in A for which $P(B) > P(A) - \varepsilon$. If we substitute P for m in Definition 2.1 we obtain the definition of

$$\int_A f(x)P(dx).$$

When $A = R^n$ and the integral exists it is called the **expectation** of f , and often denoted by $E(f)$. In particular, if C is a subset of R^n whose indicator function $\mathbf{1}_C$ is integrable, C is called an **event** and assigned the probability measure

$$P(C) = \int_{R^n} \mathbf{1}_C(x)P(dx).$$

The fundamental theorems proved for integrals over intervals in R^n carry over with only trivial change to this integral (differentiation theorems and the Fubini theorem do not), and so does the monotone convergence lemma. It follows readily that if C_1, C_2, \dots is a countable set of events, their union is an event; and if they are pairwise disjoint, then

$$P(\bigcup_j C_j) = \sum_j P(C_j).$$

Thus our integral serves all the needs of integration with respect to distributions in finite-dimensional spaces.

However, such processes as infinite sequences of tosses of a coin occur in fairly elementary probability theory, and at a more advanced level we meet stochastic processes, or random functions. An infinite sequence of real numbers is a point of R^Z , where $Z = \{1, 2, 3, \dots\}$; a function on a set T is a point of R^T . So the probability theory of such processes calls for distributions and integration over infinite-dimensional spaces R^T (where T may be Z). If we define intervals and neighborhoods in R^T in the manner that has long been customary in topology, we find that Definition 2.1 applies to this case also.

By now it should be clear that the only thing keeping us from going on and on with the full development of the Lebesgue integration theory is the fact that we have reached the end of our program of fitting the theory into undergraduate instruction (and, perhaps, of the editor's patience). Anything that can be proved about the Lebesgue integral can be proved about this integral, because it *is* the Lebesgue integral. And for those students, such as engineers, who lack time or inclination to work through detailed proofs, we are at least asking them to believe unproved statements about an integral they know, rather than about an integral whose very definition is unfamiliar to them.

This paper is an expansion of an address given to the Louisiana-Mississippi Section of the MAA on February 18, 1972.

References

1. E. J. McShane, A Riemann-type integral that includes Lebesgue-Stieltjes, Bochner and stochastic integrals, American Mathematical Society, Memoir 88 (1969).
2. Studies in Modern Analysis, M.A.A. Studies in Mathematics, Vol. 1 (R. C. Buck, editor), 1962.

HIGHLIGHTS IN THE HISTORY OF SPECTRAL THEORY

L. A. STEEN, Saint Olaf College

Not least because such different objects as atoms, operators and algebras all possess spectra, the evolution of spectral theory is one of the most informative chapters in the history of contemporary mathematics. The central thrust of the modern spectral theorem is that certain linear operators on infinite dimensional

Lynn Steen received his MIT Ph.D under Kenneth Hoffman. Since then he has been on the staff at St. Olaf College except for a year's leave at the Mittag-Leffler Institute on a N.S.F. Fellowship. He is an Associate Editor of the MONTHLY and his publications include *Counterexamples in Topology* (with J. A. Seebach, Jr., Holt Rinehart & Winston, 1970). *Editor*.

spaces can be represented in a “diagonal” form. At the beginning of the twentieth century neither this spectral theorem nor the word “spectrum” itself had entered the mathematician’s repertoire. Thus, although it has deep roots in the past, the mathematical theory of spectra is a distinctly twentieth century phenomenon.

Today every student of mathematics encounters the spectral theorem not later than his first course in functional analysis and often as early as his first course in linear algebra. Usually he studies one specimen of the spectral theorem, plucked out of historical context and imbedded in the logical context of his particular course. Although this scheme is pedagogically efficient and logically aesthetic, it does often obscure the fact that the spectral theorem was (and perhaps still is) an evolving species. Its evolution is an outstanding example of the counterpoint between pure and applied mathematics, for while the motive force in its evolution was the attempt to provide adequate mathematical theories for various physical phenomena, the forms through which it evolved are precisely those which have marked the development of modern abstract analysis.

So we offer here an austere outline of the evolution of the spectral theorem as a microcosmic example of the history of twentieth century mathematics. To understand the significance of contemporary achievements and to recognize their continuity with the past, we begin with the principal historical roots of our subject.

1. Principal axes theorem. The only theorem available at the turn of the twentieth century which we can with hindsight recognize as a direct forerunner of the modern spectral theorem is the principal axes theorem of analytical geometry. It should not be surprising that the simplest form of this theorem is contained in the writings of the founders of analytical geometry, Pierre de Fermat (1601–1665) and René Descartes (1596–1650). For the Euclidean plane R^2 , this theorem says that a quadratic form $ax^2 + 2bxy + cy^2$ can be transformed by a rotation of the plane into the normal form $\alpha x^2 + \beta y^2$, where the principal axes of the normal form coincide with the new coordinate axes. The essential content of this theorem—that the algebraic reduction to normal form corresponds to the geometric rotation onto principal axes—is contained in Descartes’ *La Géométrie* [1637], and was known at about the same time by Fermat but not published until after his death [1679]. The term “principal axes” was introduced by Leonhard Euler (1707–1783) in his investigation of the mechanics of rotating bodies [1765]; Euler also discussed (in [1748]) the reduction of quadratic forms in two and three dimensions.

The general form of the principal axes theorem asserts that any symmetric quadratic form $(Ax, x) = \sum \alpha_{ij} x_i x_j$ on R^n can be rewritten by means of an orthogonal transformation $T: R^n \rightarrow R^n$ in the normal form $\sum \lambda_i x_i^2$. (A is **symmetric** if $\alpha_{ij} = \alpha_{ji}$, and T is **orthogonal** if it leaves invariant the Euclidean metric on R^n .) The generalization from R^3 to R^n of the algebraic part of this theorem (that a quadratic form can be written as a sum of squares) was discussed by Joseph Louis Lagrange (1736–1813) in a paper [1759] on the maxima and minima of functions of several variables. In

[1827] Carl Gustav Jacob Jacobi [1804–1851] investigated the principal axes of various quadratic surfaces, and about the same time Augustin-Louis Cauchy (1789–1867) showed in [1829] and [1830] that the coefficients λ_i of the normal form of a symmetric quadratic form must be real.

But it was not until the second half of the nineteenth century that the general form of the principal axes theorem was achieved when James Joseph Sylvester (1814–1897) and Arthur Cayley (1821–1895) used the notation of matrices to systematize the algebraic description of n -dimensional space. In [1852] Sylvester showed explicitly that the coefficients λ_i in the normal form of (Ax, x) are the roots of the characteristic polynomial $\det(\lambda I - A) = 0$; in [1858] Cayley inaugurated the calculus of matrices, in which the reduction to normal form corresponded to a diagonalization process on the matrix A . Specifically, the principal axes theorem says in the language of matrices that each symmetric real matrix A is orthogonally equivalent to a diagonal matrix D ; in other words, for some orthogonal matrix T , the matrix $D = T^{-1}AT$ is in diagonal form. The diagonal entries of D are the **eigenvalues** of A , that is, the roots of the polynomial equation $\det(\lambda I - A) = 0$.

Although the new concepts of matrix theory had an immediate and profound influence on British mathematics, their impact on the continent was relatively minor. Especially in Germany bilinear forms continued well into the twentieth century to be the principal tool of analytical geometry, and in [1878] Georg Frobenius (1849–1917) published a systematic account of matrix algebra entirely in the language of bilinear forms. So by the end of the nineteenth century we can discern two versions of the principal axes theorem: the reduction to normal form of a symmetric bilinear form, and the diagonalization of a real symmetric matrix.

2. Infinite systems of linear equations. The central fact of modern spectral theory is that certain linear operators on infinite dimensional spaces can also be presented in “diagonal” form. Thus the second historical taproot of spectral theory is the evolution of infinite dimensional theory from finite dimensional cases. This evolution occurred first in algebra—in the solution of systems of linear equations—and only much later in geometry. Finite systems of linear equations were solved most often throughout the eighteenth and nineteenth centuries by the method of elimination, as expounded, for instance, in [1770] and [1779] by Euler and Etienne Bézout (1730–1783). In [1750] Gabriel Cramer (1704–1752) introduced for 3×3 systems the rule which now bears his name, although he did not, of course, use the concept or notation of determinants.

Infinite systems of equations were used throughout the eighteenth and nineteenth centuries to obtain formal solutions to differential equations by the method of undetermined coefficients: if a formal power series with unknown coefficients is substituted for the unknown in a given differential equation, the task of solving the differential equation is reduced to that of determining the infinitely many unknown coefficients. (Of course few at that time worried very much about the convergence of

the power series thus obtained.) If all went well, the infinite system of equations in the unknown coefficients would exhibit a recursive pattern which made it possible to solve the infinite system by finite dimensional tools. But for this reason precisely, these recursive techniques contributed little to the development of a general theory of infinite dimensional systems.

Joseph Fourier (1768–1830) launched the first significant general attack on the problem of infinite systems of equations when he attempted to show [1822] that every function can be expressed as an infinite linear combination of trigonometric terms. The problem of determining the unknown coefficients in these linear combinations led him directly to the general problem of solving an infinite system of linear equations. Fourier's approach (called the *principe des réduites* by Frédéric Riesz [1913a]) was to solve the first $n \times n$ system by ordinary means and let $n \rightarrow \infty$.

Although Fourier's assertion about the expansion of "arbitrary" functions into trigonometric series stimulated intense work on the theory of integration, his method of solving infinite systems of linear equations was virtually ignored. More than fifty years passed before Theodor Kötteritzsch of Saxony reopened the investigation with a paper [1870] in which he attempted to extend Cramer's rule to infinite systems. Seven years later the American astronomer George William Hill (1838–1914) published in Cambridge, Massachusetts, a monograph [1877b] in which he successfully applied to the infinite dimensional case the theory of determinants which had at that time only been established for finite dimensional systems. Hill's work was first disseminated in Europe in [1886a] when G. Mittag-Leffler reprinted it in *Acta Mathematica* in the year following the appearance in France of a paper [1885a] by Paul Appell (1855–1930) in which he applied the *principe des réduites* to determine the coefficients of the power series expansion of elliptic functions.

At this point Henri Poincaré (1854–1912) entered the discussion with two papers ([1885b], [1886b]) in which he provided a rigorous definition for an infinite determinant in order to clarify the works of Hill and Appell. The work begun in Paris by Poincaré was continued in Stockholm by Helge von Koch (1870–1924) who developed between 1890 and 1910 an extensive theory of infinite determinants. Von Koch's first major papers on this subject appeared in [1891] and [1892]; his own survey of the field in [1910d] provides further references. The more recent survey [1968] by Michael Bernkopf includes a complete discussion of these fundamental papers.

3. Integral equations. The theory of infinite matrices and determinants might have led directly to an elementary spectral theorem if someone had generalized the diagonalization form of the principal axes theorem. But the road to spectral theory was not that straight: the first spectral theorem was achieved only after infinite determinants were applied to integral equations, thereby extending the theory from the countably to the uncountably infinite. The formal study of integral equations is usually traced back to [1823] and [1826] when the young Norwegian genius Niels Henrik Abel (1802–1829) used an integral equation to solve a generalized tautochrone

problem concerning the shape of a wire along which a frictionless bead slides under the influence of gravity. Somewhat later Joseph Liouville (1809–1882) introduced (in [1837]) the method of iteration to solve a specific type of integral equation; in [1877a] Carl Neumann (1832–1925) extended Liouville's iterative method to a more general setting while investigating a boundary value problem for harmonic functions.

Neumann's work precipitated considerable research in integral equations, especially by Poincaré in France and in Rome by Vito Volterra (1860–1940). But it was not until 1900 that the theory of integral equations became especially relevant to the history of spectral theory, for in that year the Swedish mathematician Ivar Fredholm (1866–1927), then a docent at the University of Stockholm, applied to integral equations the theory of infinite matrices and determinants as developed by his colleague von Koch. By mimicking von Koch's technique for expanding infinite determinants, Fredholm developed in [1900] his now famous "alternative" theorem concerning the solutions ϕ of the integral equation

$$(1) \quad \phi(x) + \int_0^1 K(x, y)\phi(y)dy = \psi(x), \quad (0 \leq x \leq 1).$$

Just as Daniel Bernoulli (1700–1784) nearly two centuries earlier had represented the vibrating string as the limit of n oscillating particles [1732], so Fredholm considered the integral equation (1) to be the limiting case of the corresponding linear system

$$(2) \quad \phi(x_i) + \sum_{j=1}^n K(x_i, y_j)\phi(y_j) = \psi(x_i), \quad (1 \leq i \leq n).$$

Fredholm defined a "determinant" D_K for the integral equation (1) which is the continuous analog of the classical determinant of the $n \times n$ system (2) and showed—in exact analogy to the classical theory for (2)—that the integral equation (1) has a unique solution which can be expressed as the quotient of two "determinants" whenever $D_K \neq 0$; or alternatively, if $D_K = 0$, then the transposed homogeneous equation $\phi(x) + \int_0^1 K(y, x)\phi(y)dy = 0$ has nontrivial solutions and (1) is solvable if and only if ψ is orthogonal to each of these solutions. Fredholm's major paper on this subject appeared in [1903a]; a summary of this work together with later developments is the substance of his survey article [1910e].

4. David Hilbert. Although there is very little in the papers of either von Koch or Fredholm that could be construed as a logical ancestor of the modern spectral theorem, we have discussed these developments for two particular reasons—one mathematical, the other historical. The twentieth century evolution of infinite dimensional spectral theory from the much simpler finite dimensional theory is foreshadowed by the nineteenth century development of linear equation and determinant theory, from the finite to the infinite (von Koch) to the continuous (Fredholm). But there is even a more direct connection, for when Fredholm's ideas were introduced (by Fredholm's colleague Eric Holmgren) in David Hilbert's 1900–01 seminar at

Göttingen, Hilbert, in the words of Hermann Weyl [1944], “caught fire at once”. For the next ten years Hilbert (1862–1943) focused his impressive mathematical talent exclusively on integral equations, and through a series of six papers published in *Göttingen Nachrichten* from 1904 to 1910 (collected and published as one volume in [1912a]) he outlined the basic definitions and theorems of spectral theory (which he named) and Hilbert space theory (which he did not name, or even define directly).

Hilbert worked primarily with the integral equation

$$(3) \quad \phi(x) - \lambda \int_0^1 K(x, y) \phi(y) dy = \psi(x)$$

together with the analogous finite or infinite dimensional matrix equation

$$(4) \quad \phi(x_i) - \lambda \sum_j K(x_i, y_j) \phi(y_j) = \psi(x_i).$$

In the process of constructing the machinery necessary to solve these equations, Hilbert defined the spectrum of the quadratic form K , distinguished the point spectrum from the continuous spectrum, and defined the concept of complete continuity which served to separate those forms that had pure point spectra from those with more complicated spectra. But most important from the viewpoint of this essay, he formulated and proved the spectral theorem—not only for completely continuous forms, but for bounded forms as well.

Hilbert’s papers on integral equations contain an astonishing quantity of what we now recognize as modern analysis in classical language. Because he was primarily concerned with solving integral equations, Hilbert never applied his results specifically to matrices or operators; furthermore, because of the position of the parameter λ in equation (3), all of Hilbert’s eigenvalues and spectral points are reciprocals of those in use today. And while his theorems had a most modern thrust, his basic method of proof was that of Bernoulli and Fredholm—a laborious passage to the limit from the corresponding finite case.

Beginning in 1905 with his doctoral dissertation under Hilbert, Erhard Schmidt (1876–1959) generalized and simplified Hilbert’s work by introducing the suggestive language of Euclidean geometry. In [1907a], [1907b] and [1908a] Schmidt presented a definitive theory of “Hilbert’s space”—what we now call l^2 , the space of square summable sequences—replete with the language of norms, linearity, subspaces and orthogonal projections. (It was Schmidt who generalized to l^2 the iterative algorithm for orthonormalization first introduced in [1883] by Jörgen Pederson Gram of Copenhagen.) Schmidt’s conceptual simplifications were immediately incorporated by Ernst Hellinger (1883–1950) and Hermann Weyl (1885–1955) in their 1907 and 1908 dissertations under Hilbert. In [1909a] Hellinger reformulated the theory of quadratic forms in the new language of Hilbert and Schmidt, and in the same year Weyl published an extensive study of bounded forms and their spectra [1909d]. So

by the end of the first decade of the twentieth century we can perceive in the writings of Hilbert and his pupils the major part of spectral theory for bounded linear transformation on l^2 .

5. Hilbert-Schmidt spectral theory. Recall, as did Hilbert at the beginning of his first paper on integral equations [1904a], the principal axes theorem for finite dimensional spaces. Let $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ be the n (real) eigenvalues of the symmetric $n \times n$ matrix K , listed according to multiplicity. Let $\phi_1 \dots \phi_n$ be an orthonormal collection of corresponding eigenvectors, so $K\phi_i = \lambda_i\phi_i$ for $1 \leq i \leq n$. Then the action of K is represented, with respect to the basis $\phi_1 \dots \phi_n$, by the diagonal matrix L with entries λ_i on the main diagonal. The matrix T whose rows are the vectors $\phi_1 \dots \phi_n$ is an orthogonal transformation which maps the new basis vectors $\phi_1 \dots \phi_n$ back to the original (canonical) basis vectors. Thus $L = T^{-1}KT$ is the diagonalization of K by the orthogonal transformation T . The matrix L can be written as $\sum_{i=1}^n \lambda_i P_i$ where P_i is the **projection** (i.e., the transformation which projects R^n) onto the one dimensional subspace spanned by ϕ_i .

Hilbert's first step in extending this theorem was to generalize the concept of eigenvalue to the case of an infinite symmetric form K . His new concept was the **spectrum** of K , denoted by $\sigma(K)$, which is the set of λ for which the transformation $\lambda I - K$ is not invertible. (Actually Hilbert used $I - \lambda K$ while Schmidt used $\lambda I - K$.) The subset of $\sigma(K)$ consisting of those λ for which the equation $K\phi = \lambda\phi$ has non-trivial solutions is called the **point spectrum** of K ; this is the strict analog of the set of eigenvalues. The complement of the point spectrum in $\sigma(K)$ is called the **continuous spectrum**. Much of Hilbert's fourth paper [1906a] is devoted to a study of the relationships between a transformation K and its spectrum $\sigma(K)$.

One of the simplest relationships Hilbert discovered was that the spectrum of K is a bounded set whenever K is a bounded transformation—that is, whenever the set $S = \{\|Kx\| : \|x\| \leq 1\}$ is bounded, where the notation $\|\ \|$, due to Schmidt, is the l^2 norm. In fact, whenever K is symmetric, the least upper bound of S , called the bound (or norm) of K and denoted by $\|K\|$, is the same as the least upper bound of $\{|\lambda| : \lambda \in \sigma(K)\}$; this fact is now called the spectral radius theorem. The bounded linear transformations on l^2 are important from another point of view, also due to Hilbert: they are precisely the continuous linear transformations, in the sense that they preserve strong convergence (i.e., $\|Kx_n - Kx\| \rightarrow 0$ whenever $\|x_n - x\| \rightarrow 0$).

Hilbert extended the principal axes theorem to symmetric bounded linear transformations; the spectra of these transformations are bounded subsets of the real axis. Those λ in the point spectrum $\rho(K)$ of K are like eigenvalues since there exists an orthonormal collection of corresponding eigenvectors ϕ_λ satisfying $K\phi_\lambda = \lambda\phi_\lambda$. If P_λ denotes the projection onto the subspace generated by ϕ_λ , we can form the diagonal transformation $L = \sum \lambda P_\lambda$ where λ ranges over the point spectrum $\rho(K)$. The transformation L reflects accurately the action of K on the subspace generated by the eigenvectors ϕ_λ , but since this subspace will in general be strictly smaller than

l^2 —since we have omitted the continuous spectrum—we cannot say that L and K represent the same transformation.

To express the contribution of the continuous spectrum, Hilbert set up an integral patterned after one defined in [1894] by the Dutch mathematician Thomas-Jean Stieltjes (1856–1894). In his study of continued fractions, Stieltjes was led (via the problem of moments) to the integral $\int_a^b f(x) dg(x)$ as the limit of the sum $\sum f(\xi_i) [g(x_i) - g(x_{i-1})]$ (for continuous f and increasing g). By rewriting the sum $\sum \lambda_i P_{\lambda_i}$ as $\sum \lambda_i [E_{\lambda_i} - E_{\lambda_{i-1}}]$, where $E_{\lambda_i} = \sum_{j=1}^i P_{\lambda_j}$, Hilbert constructed for the continuous spectrum $s(K)$ the Stieltjes-type integral $\int_{s(K)} \lambda dE_\lambda$ as the limit of sums of the form $\sum \lambda_i [E_{\lambda_i} - E_{\lambda_{i-1}}]$. Then Hilbert's spectral theorem was that every symmetric bounded linear transformation on l^2 can be represented (by means of an orthogonal transformation) in the “diagonal” form

$$(5) \quad \sum_{\rho(K)} \lambda P_\lambda + \int_{s(K)} \lambda dE_\lambda$$

where the summation is over the point spectrum, and the integral is over the continuous spectrum.

Hilbert completed his spectral theory by identifying a large class of transformations whose continuous spectra were empty. He called these transformations **completely continuous**, and Schmidt characterized them by the property of mapping weakly convergent sequences to strongly convergent sequences. In other words, the linear transformation K is completely continuous if

$$\|Kx_n - Kx\| \rightarrow 0$$

whenever $(y, x_n) \rightarrow (y, x)$ for all y . The completely continuous transformations are the nearest infinite dimensional analog to the finite dimensional transformations, since their spectra consist entirely of eigenvalues with zero as the only possible accumulation point; furthermore, every completely continuous symmetric linear transformation K can be expressed (by an orthogonal transformation) in the diagonal form $\sum \lambda P_\lambda$ (since $s(K) = \phi$).

Although Hilbert originally used infinite matrices merely as convenient approximations to integral equations, he concluded his theoretical investigation by establishing a major link between these two theories, namely that of a **complete orthogonal system**. Such a system $\{\phi_n\}$, either of vectors in the sequence space l^2 or of continuous functions on the interval $[0, 1]$, is characterized by the orthogonality relation $(\phi_n, \phi_m) = 0$ if $n \neq m$, together with the fact that every vector (or continuous function) ϕ can be represented by the Fourier-type series $\phi = \sum_{n=1}^{\infty} a_n \phi_n$. The matrix equation (4) can then be derived (by mathematics, rather than by analogy) from the integral equation (3) by replacing each continuous function ϕ, ψ by its Fourier expansion with respect to the complete orthogonal system $\{\phi_n\}$. This application of a complete orthogonal system enabled Hilbert to derive Fredholm's alternative

theorem for the integral equation (1) directly from the corresponding theorem for the infinite linear system (2).

To keep the record straight, we should emphasize again that Hilbert introduced spectral theory in the language of quadratic forms, whereas we have reported his work primarily in the language of linear transformations on the infinite dimensional space l^2 . Barely fifty years had elapsed since Cayley in England and Hermann Grassman (1809–1877) in Germany had begun, in [1843] and [1844] the systematic study of Euclidean n -dimensional space for $n > 3$. Hilbert and Schmidt were the first to explore the totally unknown depths of an infinite dimensional space and it was not until other such spaces were studied that the broad outlines of a theory of linear transformations became clear. The early twentieth century development of infinite dimensional (function) spaces is recorded in [1966].

6. The Lebesgue Integral. At about the same time as Hilbert was creating his spectral theory for spaces of square summable sequences, Henri Lebesgue (1875–1941) was developing the new integral which now bears his name ([1901], [1904c]). In three brief papers in 1907 Friedrich Riesz (1880–1956) and Ernst Fischer (1875–1959) joined together the works of Hilbert and Lebesgue by showing that Hilbert's space l^2 is isomorphic to the space L^2 of functions whose square is Lebesgue integrable. In a subsequent paper [1910c] (in which he introduced the more general L^p spaces), Riesz derived a spectral theory for L^2 entirely analogous to that developed for l^2 by Hilbert and Schmidt.

In the year preceding the appearance of his paper on L^p spaces, Riesz proved in [1909b] his now famous representation theorem in which he solved a problem first studied in [1903b] by Jacques Hadamard (1865–1963). What Riesz showed was that every continuous linear functional on $C([a, b])$ is a Stieltjes integral $\int f dg$ with respect to some function g of bounded variation. Lebesgue then showed in [1910f], in direct response to Riesz's paper, that every Stieltjes integral can be interpreted as a Lebesgue integral under a proper interpretation of the heuristic formula

$$\int f(x) dg(x) = \int f(x) g'(x) dx.$$

This led Johann Radon (1887–1956) to develop (in [1913b]) integration with respect to a measure (i.e., a countably additive set function) thus encompassing the integrals of both Lebesgue and Stieltjes and providing the foundation for all modern theories of the abstract integral.

We can see from this digression that the evolution of the modern integral was closely connected to Hilbert's creation of spectral theory. Although neither theory depended logically on the other, the historical dependence of each on the other is quite clear: Hilbert used Stieltjes' integral to obtain the spectral theorem for l^2 , while Riesz, following Hilbert, used and thereby immortalized Lebesgue's integral by developing the spectral theory of L^2 .

The second decade of spectral theory was rather uneventful. In Göttingen, Hilbert had turned his attention to the axiomatization of physics, a task which he had proposed to the International Congress of Mathematicians in 1900 as the sixth of his famous 23 problems for twentieth century mathematics. "Physics," he said, "is much too hard for physicists" ([1970]). In the United States Eliakim Hastings Moore (1862–1932) at the University of Chicago developed a system of "general analysis" ([1908b], [1912b]) which was designed to include as special cases the work of Hilbert, Fredholm and Riesz. But Moore's results were constrained by the fact that European investigators were not then accustomed to receiving new mathematical ideas from America. So while Moore's research had a profound effect on the development of mathematics in the United States, it did not influence significantly the direction of research on spectral theory.

Many European efforts from 1910 to 1925 were devoted to exposition and recapitulation. Riesz [1913a], Fredholm [1910e] and von Koch [1910d] published surveys of the theory of infinitely many variables and integral equations, each of which contained various forms of Hilbert's spectral theory. Hilbert's collected papers on integral equations were themselves published in book form in [1912a]. But certainly the most impressive survey work of this period was the massive *Enzyklopädie der Mathematischen Wissenschaften* which contains in volume II.3.2. a comprehensive discussion of integral equations and spectral theory by Hellinger and Otto Toeplitz (1881–1940); this survey paper was also published separately [1928a].

7. Quantum mechanics. In Göttingen in 1925–26 Werner Heisenberg (1901–) and Erwin Schrödinger (1887–1961) created the theory of quantum mechanics. In Heisenberg's theory the physical fact that certain atomic observations cannot be made simultaneously was interpreted mathematically to mean that the operations which represented these observations were not commutative. Since the algebra of matrices is non-commutative, Heisenberg together with Max Born and Pascual Jordan ([1925a], [1926a]) represented each physical quantity by an appropriate (finite or infinite) matrix, called a transformation; the set of possible values of the physical quantity was the spectrum of the transformation. (So the spectrum of the transformation which represented the energy of an atom was precisely the spectrum of the atom.)

Schrödinger, in contrast, advanced a less unorthodox theory based on his partial differential wave equation. Following some initial surprise that Schrödinger's "wave mechanics" and Heisenberg's "matrix mechanics"—two theories with substantially different hypotheses—should yield the same results, Schrödinger unified the two approaches by showing, in effect, that the eigenvalues (or more generally, the spectrum) of the differential operator in Schrödinger's wave equation determine the corresponding Heisenberg matrix. Similar results were obtained simultaneously ([1925b], [1926b]) by the British physicist Paul A. M. Dirac (1902–). Thus interest in spectral theory once again became quite intense.

Hilbert himself was astonished that the spectra of his quadratic forms should come to be interpreted as atomic spectra. "I developed my theory of infinitely many variables from purely mathematical interests, and even called it 'spectral analysis' without any presentiment that it would later find an application to the actual spectrum of physics" [1970]. It quickly became clear, however, that Hilbert's spectral theory was the proper mathematical basis for the new mechanics. Finite and infinite matrices were interpreted as transformations on a Hilbert space (still thought of primarily as l^2 or L^2) and physical quantities were represented by these transformations. The mathematical machinery of quantum mechanics became that of spectral analysis and the renewed activity precipitated the publication by Aurel Wintner (1903–1958) of the first book [1929b] devoted to spectral theory.

Hilbert's original spectral theorem applied to real quadratic forms (or infinite matrices) that were bounded and symmetric. This theorem was quickly and easily extended (by Schmidt and others) to bounded complex matrices $A = (a_{ij})$ for which $a_{ij} = \bar{a}_{ji}$; such matrices are called **Hermitian** after the French mathematician Charles Hermite (1822–1901) who introduced them (in [1855]) and proved their eigenvalues real. Both symmetric and Hermitian forms may be characterized in terms of their respective inner product by the relation $(Ax, y) = (x, Ay)$ for all x, y . Like symmetric matrices, Hermitian transformations have real spectra and, more generally, play the role of the real number line in the algebra of transformations.

Almost miraculously, it was precisely the Hermitian transformations which qualified in the new mechanics to represent a physical quantity. One reason for this is that physical quantities are measured by real numbers, so it is natural to represent them by those transformations which behave like real numbers. Perhaps a more compelling justification is that the hypothesis that the transformations of mathematical physics are Hermitian implies certain fundamental laws (or assumptions) of physics: if A is Hermitian, the wave equation $\dot{\phi} = A\phi$ implies the conservation of energy, a fundamental law of classical mechanics, and the solutions of Schrödinger's equation $\dot{\phi} = iA\phi$ will have constant norm, which is a fundamental assumption of quantum mechanics.

Although every observable was represented in the new mechanics by a Hermitian transformation, it was not necessarily true that every such transformation represented an observable. Dirac [1930b] added the crucial hypothesis that a Hermitian transformation represents an observable if and only if its eigenvectors form a complete (orthogonal) system: his hypothesis was designed to insure that any vector (representing a quantum mechanical state) could be expressed as a (possibly infinite) linear combination of eigenvectors of any given observable. The identification of transformations with this property is part of the Hilbert-Schmidt spectral theory, but this theory provided only a partial answer: those Hermitian transformations which are completely continuous have a complete set of eigenvalues.

This theorem did not provide a satisfactory elucidation of Dirac's hypotheses

since the transformations of quantum mechanics are usually not completely continuous. Most of the important transformations in physics involve differentiation of, say, functions in L^2 . The theorem on integration by parts shows that differentiation is formally symmetric, for in this case $(Af, g) = (f, Ag)$ means $\int f'g = \int fg'$. But since the derivative of a function has practically no relation to the magnitude of the function, differentiation is neither continuous nor bounded, nor even defined everywhere. In fact, if a symmetric or Hermitian transformation (like differentiation) were defined everywhere, it would have to be bounded. This rather surprising result—which says, in effect, that a candidate for the spectral theorem which fails to be bounded must fail to be everywhere defined—was demonstrated as early as [1910b] by Hellinger and Toeplitz.

Thus many of the transformations of quantum mechanics, although Hermitian, failed nevertheless to satisfy the second of Hilbert's hypotheses, namely, that they be bounded. Like differentiation, they were unbounded and defined only on a dense subset of L^2 . Paul Dirac attempted to overcome the exceptional behavior of differentiation by introducing his δ -function to provide derivatives where none existed and thereby to enlarge the set of functions to which the differentiation transformation could be applied. Dirac's approach was highly successful in explaining the new quantum mechanics and led eventually to Laurent Schwartz' theory of distributions precisely because it lacked an adequate mathematical foundation. But in 1926 Dirac's approach represented more an alternative to spectral theory than an extension of it, and it did not really help to extend Hilbert's theory to unbounded transformations.

8. John von Neumann. After Hilbert, the only major study of unbounded transformations was that published in [1923] by Torsten Carleman (1892–1949) in Sweden. In this monograph Carleman showed that many of the results of Fredholm and Hilbert still hold under a weaker type of boundedness hypothesis. But from the viewpoint of spectral theory, the major breakthrough came in 1927–29 when the twenty-five year old Hungarian John von Neumann (1903–1957) revolutionized the study of spectral theory by introducing the abstract concept of a linear operator on Hilbert space. In [1927] von Neumann expressed the transformation theory of quantum mechanics in terms of operators on a Hilbert space, and explicitly recognized the need to extend from the bounded to the unbounded case the spectral theory of Hermitian operators. In [1929a] he carried out that extension.

Before von Neumann, the name "Hilbert space" had been applied principally to the space l^2 of square summable sequences (often called "Hilbert's space") or to the space L^2 of Lebesgue square integrable functions which Riesz had proved isomorphic to l^2 . The essential properties of these spaces, widely recognized, were those of a vector space with an inner product which was complete and separable (i.e., which had a countable dense subset). Von Neumann's first step in his theory of linear operators was to define an (abstract) Hilbert space axiomatically as any

separable, complete inner product space. He then defined a **general linear operator** on the abstract Hilbert space H as a linear transformation defined on some subset of H . This subset, called the domain D_T of the operator T , is usually assumed to be a linear subspace of H , which, like the domain of the differentiation operator, is dense in H . Von Neumann's linear operators thus comprehend both the matrices and quadratic forms of Hilbert's theory, and the transformations of quantum mechanics.

A linear operator is continuous if and only if it is bounded, and a bounded linear operator with a dense domain can be uniquely extended to a bounded linear operator on the whole space H . Every linear operator T with a dense domain has a unique **adjoint operator** T^* defined by the relation

$$(6) \quad (Tx, y) = (x, T^*y)$$

for all $x \in D_T$; the domain of T^* is the set of $y \in H$ for which (6) holds for all x . An operator T is called **self-adjoint** if $T = T^*$, and **symmetric** if T^* is an extension of T , or equivalently, if $(Tx, y) = (x, Ty)$ whenever $x, y \in D_T$. (In von Neumann's papers, the self-adjoint operators were called **hypermaximal**.)

Every self-adjoint operator is clearly symmetric, and every symmetric operator which is everywhere defined must be self-adjoint. Thus for bounded linear operators (which either are everywhere defined or may be extended to become so) the concept of symmetric and self-adjoint coincide. The Hellinger-Toeplitz theorem, cited in section 7 above, can be extended to von Neumann's operators and shows that any symmetric operator which is everywhere defined must be bounded. (This result is closely related to a more general theorem due to Stefan Banach (1892–1945), now commonly known as the closed graph theorem [1932b].) Thus in von Neumann's theory there are precisely three types of symmetric operators:

- I. bounded, self-adjoint and everywhere defined;
- II. unbounded, self-adjoint and densely but not everywhere defined; and
- III. unbounded, not self-adjoint, and densely but not everywhere defined.

Hilbert's original theory applied to operators of type I, while von Neumann's spectral theorem encompassed those of type II as well since it applies to all self-adjoint operators. This theory, though initiated by von Neumann, was developed by Riesz [1930c] and more extensively, by Marshall H. Stone (1903–) at Yale University who expounded it in great detail in [1932a]. The combined (but largely independent) efforts of von Neumann and Stone for the five year period 1927–1932 provided for spectral theory the largest collection of new methods since Hilbert's five year effort of 1901–1906.

9. Von Neumann — Stone Spectral Theory. Hilbert's general spectral theorem says that every bounded symmetric linear transformation T can be written in the form

$$\sum_{\rho(T)} \lambda P_\lambda + \int_{s(T)} \lambda dE_\lambda.$$

By rewriting the first sum as a Stieltjes-type integral and combining it with the second integral, we may express Hilbert's spectral theorem in the concise form

$$(7) \quad T = \int_{\sigma(T)} \lambda dE_\lambda,$$

where the integral is over the entire (bounded) spectrum of T . The operators E_λ are projections with the following properties:

- (i) If $\lambda < \mu$, the range of E_λ is contained in the range of E_μ ;
- (ii) If $\varepsilon > 0$, $E_{\lambda+\varepsilon} \rightarrow E_\lambda$ as $\varepsilon \rightarrow 0$;
- (iii) $E_\lambda \rightarrow 0$ as $\lambda \rightarrow -\infty$;
- (iv) $E_\lambda \rightarrow I$ as $\lambda \rightarrow +\infty$.

Stone called such a family of operators a **resolution of the identity**; in more intuitive language, properties (i)–(iv) require that the function $\lambda \rightarrow E_\lambda$ be increasing, continuous from the right, with 0 and I as left and right limiting values.

The von Neumann-Stone extension of the spectral theorem for self-adjoint operators from the bounded to the unbounded case corresponds to the extension of (7) from bounded to unbounded spectra $\sigma(T)$. Specifically, it says that to each self-adjoint operator T there corresponds a unique resolution of the identity $\{E_\lambda\}$ such that (7) holds.

Despite the power of this theorem, many differential operators are not covered by it since they are rarely self-adjoint. For instance, to make the operator $D = d/dt$ symmetric on a dense subset Δ of the Hilbert space $L^2(0, 1)$, we should select for Δ the subset consisting of those continuously differentiable functions f which satisfy $f(0) = f(1) = 0$ (in order to insure that the relation $(Df, g) = (f, Dg)$ would follow by integration by parts). But the domain Δ is too small to permit D to be self-adjoint, for *every* continuously differentiable L^2 function is in the domain of D^* . To make D self-adjoint we would have to enlarge its domain appropriately—thereby risking a loss of symmetry. Each symmetric operator of type III suffers from the same disease: its domain is smaller than that of its adjoint. Moreover the cure—namely, extension of the domain—is often fatal since with a larger domain the operator may fail to be symmetric.

To apply his spectral theorem to symmetric operators von Neumann had to know which types of symmetric operators admit self-adjoint extensions. He [1929a] and Wintner [1929b] identified a large class of such operators, namely those operators T , called **semibounded**, for which there is a positive constant M satisfying either $(Tx, x) \leq M \|x\|^2$ for all $x \in D_T$ or $-M \|x\|^2 \leq (Tx, x)$ for all $x \in D_T$. The best statement of this result is due to Stone [1932a] and Kurt O. Friedrichs [1934]: every semibounded symmetric operator may be extended to a semibounded self-adjoint operator with the same bound.

Whereas the central focus of the von Neumann-Stone spectral theory (and of Hilbert's also) is on operators with real spectra, the spectral theorem does apply, at

least in two cases, to operators with more general spectra. The simplest case concerns *isometric* operators which leave the inner product on H invariant; from this definition it follows easily that the spectrum of an isometric operator is a subset of the unit circle. An isometric operator that maps H onto H is called *unitary* and is characterized by the fact that its adjoint is its inverse (i.e., $TT^* = T^*T = I$). Unitary operators were first studied in [1909c] by Isaac Schur following their introduction by Léon Autonne in [1902]. In [1929a] von Neumann employed the Cayley transform $C: T \rightarrow (T - iI)(T + iI)^{-1}$ to map symmetric operators T into isometric operators $C(T)$; he showed that T is self-adjoint if and only if $C(T)$ is unitary. Thus the spectral theory for unitary operators follows from that for self-adjoint operators by use of a spectral integral on the unit circle instead of on the real line.

Now every bounded linear operator T can be written in the “Cartesian” form $T = A + iB$, where A and B are bounded and self-adjoint; in fact,

$$A = \frac{1}{2}(T + T^*), \quad B = \frac{1}{2i}(T - T^*).$$

Thus it would appear likely that the spectral theorem could be extended to all bounded linear operators by using this decomposition. However, the details of that extension require that $AB = BA$ (or equivalently, that $TT^* = T^*T$). So the desired extension works only for those operators which commute with their adjoints: such operators are called **normal**, after Toeplitz [1918a]. Toeplitz extended Hilbert’s spectral theorem to completely continuous normal quadratic forms by showing that such a form was unitarily equivalent to a diagonal form. More generally, the spectral resolution

$$T = \int_{\sigma(T)} \lambda \, dE_\lambda$$

extends to bounded normal operators, where the integration is over the spectrum of T which is a compact subset of the complex plane contained in the disc of radius $\|T\|$. Von Neumann [1930a] and Stone [1932a] extended both the definition and spectral theory of normal operators to the unbounded case as well.

We have come a long way from the principal axes theorem, and the spectral theorems of von Neumann and Stone reflect far more analysis than geometry. The geometric content of the spectral theorem for finite dimensional space is that the entire space can be expressed as the direct sum of subspaces on each of which the given transformation acts like simple multiplication. But this theorem fails in the infinite dimensional cases as soon as the continuous spectrum appears. In a paper written in 1938 but not published until [1949a], von Neumann effectively resuscitated the geometrical spectral theorem by defining a direct integral of Hilbert spaces (in strict analogy with the direct sum). He then showed that the action of a self-adjoint operator on any Hilbert space could be represented as the accumulated effect of

simple multiplications on certain subspaces whose direct integral was (unitarily equivalent to) the original space.

10. Gelfand-Naimark Theorem. The collection of all operators on a Hilbert space forms a ring; such rings, with various topologies, were extensively investigated by von Neumann and Francis J. Murray in [1936a], [1937a] and [1940a]. During the same period 1936–40 S. W. P. Steen published in England a series of five papers ([1936b], [1937b], [1938a], [1939], [1940b]) devoted to an axiomatic theory of operators. But the papers that offered the most significant insight into the spectral theorem were [1941a], [1941b] and [1943] published in the U.S.S.R. by Israel M. Gelfand, Mark A. Naimark and Georgii E. Silov. Gelfand and his colleagues created a theory of normed rings which not only subsumed much of the work of von Neumann, Murray and Steen on rings of operators, but also provided a beautiful general setting for the study of Fourier transforms and harmonic analysis. Related studies were carried out in the United States by Stone ([1940c], [1941c]) and Shizuo Kakutani [1941d].

Normed rings were first introduced in [1936c] by the Japanese mathematician. Mitio Nagumo under the name of “linear metric rings”. In [1946] Charles E. Rickart christened Gelfand’s normed rings “Banach algebras” to avoid misunderstanding due to the algebraic meaning of “ring”; as a consequence Russian mathematicians now use the former name, while Americans use the latter. But regardless of its name, the properties of a **Banach algebra** are those of a complete normed algebra (over the complex field C) satisfying the multiplicative triangle inequality $\|x\| \|y\| \leq \|xy\|$. We shall assume that each Banach algebra contains an identity element e , where $\|e\| = 1$. The set of all bounded linear operators on a Hilbert space is a Banach algebra, as is the set of all continuous complex-valued functions on a compact topological space X (with the sup norm $\|f\| = \sup \{|f(x)| : x \in X\}$). The part of Banach algebra theory germane to spectral theory is the relation between these two examples.

Gelfand’s theory of commutative Banach algebras depends on three fundamental concepts: homomorphisms, maximal ideals and spectra. A *homomorphism* of a commutative Banach algebra B is a non-zero multiplicative linear functional; its kernel is a **maximal ideal** since it is contained in no larger proper ideal. Moreover, every maximal ideal I is the kernel of some homomorphism for in this case the factor algebra B/I is the field C of complex numbers (according to a result announced by Stanislaw Mazur [1938b] and proved by Gelfand [1941a]) so the composite map $B \rightarrow B/I \rightarrow C$ is a homomorphism of B whose kernel is I . The set M_B of homomorphisms (or equivalently, of maximal ideals) of B is given the weakest topology relative to which all of the functions $\hat{x}: h \rightarrow h(x)$ are continuous, for all $x \in B$. Then the topological space M_B is compact and Hausdorff, and each element x of B is represented in $C(M_B)$ (the Banach algebra of continuous complex valued functions on M_B) by its “Gelfand transform” \hat{x} .

In strict analogy with the spectral theory of operators on a Hilbert space, Gelfand defines the *spectrum* $\sigma(x)$ of an element $x \in B$ to be the set of complex numbers λ for which the element $x - \lambda e$ has no inverse. The set $\sigma(x)$ is compact, non-empty and contained in the disc of radius $\|x\|$. Furthermore, $\sigma(x)$ happens to be precisely equal to the range of the Gelfand transform \hat{x} : $\sigma(x) = \{h(x) \mid h \in M_B\}$. For this reason the space M_B of maximal ideals is often called the *spectrum* of the Banach algebra B . If B is the algebra generated by a single element x (such as a particular operator on H), then the spectrum of the algebra B is mapped homeomorphically by \hat{x} onto the spectrum of x .

In [1943] Gelfand and Naimark showed that the commutative Banach algebra $C(M_B)$ is characterized by the presence of an *involution*, namely the operation of complex conjugation $*$: $f \rightarrow \bar{f}$. Specifically, they showed that any commutative Banach algebra with an involution (called a *B*-algebra*) is isometrically isomorphic to the algebra $C(M_B)$ for some Banach algebra B . In particular, the commutative *B** algebra $B(T)$ generated by a given bounded normal operator T is isomorphic to the algebra $C(\sigma(T))$ of all continuous functions on $\sigma(T)$, the spectrum of $B(T)$; T is assumed normal in order that the presence of the involution $*$: $T \rightarrow T^*$ should not destroy the commutativity of the algebra $B(T)$.

The impact of the Gelfand-Naimark theorem on spectral theory is this: the spectral theorem for a bounded normal operator T can be inferred via the isomorphism between $B(T)$ and $C(\sigma(T))$ from a corresponding theorem concerning continuous complex valued functions on $\sigma(T)$. The required theorem is just that every continuous function f on $\sigma(T)$ (in particular, the identity function $f(\lambda) = \lambda$) can be approximated uniformly by measurable step functions of the form $\sum f(\lambda_i) \chi_{\Lambda_i}$, where χ_{Λ_i} is the characteristic function of the measurable set Λ_i . The translation of this theorem to the algebra $B(T)$ (in the special case $f(\lambda) = \lambda$) is the spectral resolution of the bounded normal operator T : $T = \int \lambda dE_\lambda$. In words instead of symbols, the approximation theorem says that a continuous function can be approximated by linear combinations of characteristic functions, while the spectral theorem says that bounded normal operators can be approximated by linear combinations of projections. Thus Gelfand's theory of Banach algebras revealed that the spectral theorem is in some fundamental sense equivalent to a most rudimentary fact in the theory of functions.

Gelfand's theory actually yields a spectral theorem far stronger than those which we have so far discussed. By translating the approximation theorem for an arbitrary continuous function f we obtain a spectral resolution of the form $f(T) = \int f(\lambda) dE_\lambda$. This formula was originally introduced by von Neumann and Stone as the basis of their "operational calculus". A related general spectral theorem, also due to von Neumann [1930a], can be inferred from the Gelfand-Naimark isomorphism: any commutative family of bounded normal operators admits a simultaneous diagonalization—that is, a single resolution of the identity which simultaneously represents all operators in the family by means of the integral $\int f(\lambda) dE_\lambda$ for various functions f .

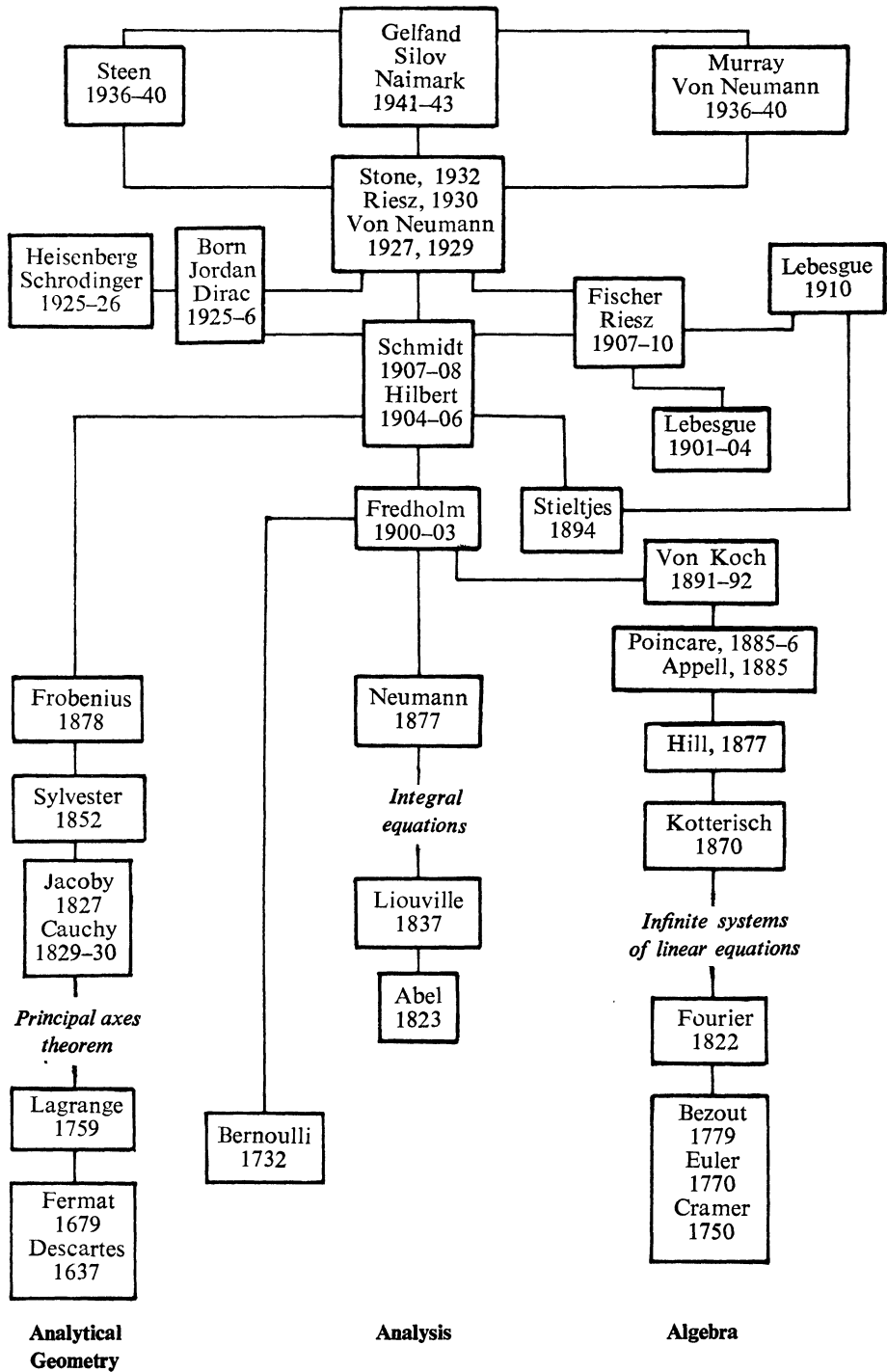


FIG. 1

11. Unfinished business. This concludes our saga of the spectral theorem. Our historical vision has been deliberately narrow, focused throughout on the evolution of just one theorem and only rarely have we glanced at the many fascinating applications and extensions of the basic theory. For example, spectral theory for spaces without inner products can be traced back to Riesz [1918b] and T.H. Hildebrandt [1931], while the rudiments of spectral theory for differential operators are contained in the work [1908c] of George Birkhoff; in [1928b] and [1930d] Norbert Wiener developed a theory of **spectral analysis** for functions in an attempt to analyze mathematically the spectrum of white light, while twenty years later Arne Beurling [1949b] inaugurated the complementary study of *spectral synthesis*; and in [1942] Edgar R. Lorch, continuing work begun in [1913a] by F. Riesz, investigated spectral sets in the plane by means of contour integrals.

Had we stopped to investigate each such offshoot our evolutionary tree (Figure 1) would have looked like a forest. Indeed, it took Nelson Dunford and Jacob Schwartz nearly 3000 pages to survey spectral theory ([1958a], [1963], [1971]). So any who are inspired to examine the fruits of spectral theory are invited to read this treatise or any of its many less ambitious companions ([1951], [1953], [1958b], [1962]). Our mission to describe the roots and main trunk of spectral theory is accomplished.

References

- 1637 Rene Descartes, La Géométrie, Appendix I to Discours de la Méthode, Leiden, 1637, (Oeuvres, VI, 367–514).
- 1679 Pierre de Fermat, Ad locus planos et solidos isagoge, Varia Opera Mathematica, 1679, (Oeuvres, I, 91–110).
- 1732 Daniel Bernoulli, Theoremata de oscillationibus corporum filo flexili connaxorum et catenae verticaliter suspensae, Comm. Acad. Scient. Imper. Petropolitanae, 6 (1732) 108–122.
- 1748 Leonhard Euler, Introductio in Analysin Infinitorum, 1748, (Opera Omnia (1), IX, 379–392).
- 1750 Gabriel Cramer, Introduction à l'analyse des lignes courbes algébriques, Geneva, 1750.
- 1759 Joseph-Louis Lagrange, Recherches sur la méthode de maximis et minimis, Miscellanea Taurinensia, 1 (1759) 18–42, (Oeuvres, I, 3–20).
- 1765 Leonhard Euler, Theoria Motus Corporum Solidorum Seu Rigidorum, 1765, (Opera Omnia (2) III, 193–214).
- 1770 Leonhard Euler, Vollständige Ableitung zur Algebra, St. Petersburg, 1770.
- 1779 Etienne Bézout, Théorie générale des équations algébriques, Paris, 1779.
- 1822 Joseph B. J. Fourier, Théorie analytique de la chaleur, Paris, 1822.
- 1823 Niels Henrik Abel, Solution de quelques problèmes à l'aide d'intégrales définies, Magazin for Naturv., 1 (1823) 205–215, (Oeuvres, I, 11–27).
- 1826 Niels Henrik Abel, Résolution d'un problème de mécanique, J. Reine Angew. Math., 1 (1826) 153–157, (Oeuvres, I, 97–101).
- 1827 Carl G. J. Jacobi, Über die Hauptaxen der Flächen der Zweiten Ordnung, J. Reine Angew. Math., 2 (1827) 227–233, (Werke, III, 45–53).
- 1829 Augustin-Louis Cauchy, Sur l'équation à l'aide de laquelle on détermine les inégalités séculaires des mouvements des planètes, Exercices de Mathématiques, Paris, 1829, (Oeuvres (2), IX, 174–195).

- 1830 Augustin-Louis Cauchy, *Mémoire sur l'équation qui a pour racines les moments d'inertie principaux d'un corps solide et sur diverses équations du même genre*, *Mém. Acad. Sci. Inst. France*, 9(1830) 111–113, (*Oeuvres*, (I), II, 79–81).
- 1837 Joseph Liouville, *Sur le développement des fonctions...* (second mémoire), *J. Math. Pures Appl.* 2(1837) 16–35.
- 1843 Arthur Cayley, *Chapters in the analytical geometry of n dimensions*, *Camb. Math. J.*, 4(1843), 119–127, (*Math. Papers*, I, 55–62).
- 1844 Hermann Grassmann, *Die Ausdehnungslehre*, Leipzig, 1844.
- 1852 James Joseph Sylvester, *A demonstration of the theorem that every homogeneous quadratic polynomial is reducible by real orthogonal substitution to the form of a sum of positive and negative squares*, *Phil. Mag.*, 4(1852) 138–142, (*Math. Papers*, I, 378–381).
- 1855 Charles Hermite, *Remarque sur un théorème de M. Cauchy*, *C. R. Acad. Sci. Paris*, 41 (1855) 181–183.
- 1858 Arthur Cayley, *A memoir on the theory of matrices*, *Philos. Trans. Roy. Soc. London*, 148 (1858) 17–37, (*Math. Papers*, II, 475–496).
- 1870 Theodor Kötteritzsch, *Über die Auflösung eines Systems von unendlich vielen linearen Gleichungen*, *Z. Math. Physik*, 15(1870) 1–15, 229–268.
- 1877a Carl Neumann, *Untersuchungen über des logarithmische und Newtonsche Potential*, Teubner, Leipzig, 1877.
- 1877b George William Hill, *On the Part of the Motion of the Lunar Perigee which is a Function of the Mean Motions of the Sun and Moon*, John Wilson, Cambridge, Mass., 1877.
- 1878 Georg Frobenius, *Über lineare Substitutionen und bilineare Formen*, *J. Reine Angew. Math.*, 84 (1878) 1–63, (*Gesammelte Abhandlungen*, I, 343–405).
- 1883 Jörgen Pederson Gram, *Über die Entwicklung realer Funktionen in Reihen mittelst der Methode der kleinsten Quadrate*, *J. Reine Angew. Math.*, 94 (1883) 41–73.
- 1885a Paul Appell, *Sur une méthode élémentaire pour obtenir les développements en série trigonométrique des fonctions elliptiques*, *Bull. Soc. Math. France*, 13(1885) 13–18.
- 1885b Henri Poincaré, *Remarques sur l'emploi de la méthode précédente*, *Bull. Soc. Math. France*, 13(1885) 19–27, (*Oeuvres*, V, 85–94).
- 1886a George William Hill, *On the part of the motion of the lunar perigee which is a function of the sun and the moon*, *Acta Mathematica*, 8(1886) 1–36, (*Coll. Math. Works*, I, 243–270).
- 1886b Henri Poincaré, *Sur les déterminants d'ordre infini*, *Bull. Soc. Math. France*, 14 (1886) 77–90, (*Oeuvres*, V, 95–107).
- 1891 Helge von Koch, *Sur une application des déterminants infinis à la théorie des équations différentielles linéaires*, *Acta Mathematica*, 15(1891) 53–63.
- 1892 Helge von Koch, *Sur les déterminants infinis et les équations différentielles linéaires*, *Acta Mathematica*, 16(1892) 217–295.
- 1894 Thomas-Jean Stieltjes, *Recherches sur les fractions continues*, *Ann. Fac. Sci. Toulouse*, 8(1894) J. 1–122, (*Oeuvres*, II, 402–566).
- 1900 Ivar Fredholm, *Sur une nouvelle méthode pour la résolution du problème de Dirichlet*, *Öfver. Vet. Akad. Förhand*, Stockholm, 57 (1900) 39–46.
- 1901 Henri Lebesgue, *Sur une généralisation de l'intégrale définie*, *C. R. Acad. Sci. Paris*, 132 (1901) 1025–1028.
- 1902 Léon Antonne, *Sur l'Hermitien*, *Rend. Circ. Mat. Palermo*, 16 (1902) 104–128.
- 1903a Ivar Fredholm, *Sur une classe d'équations fonctionnelles*, *Acta Mathematica*, 27 (1903) 365–390.
- 1903b Jaques Hadamard, *Sur les opérations fonctionnelles*, *C. R. Acad. Sci. Paris*, 136 (1903) 351–354.

- 1904a David Hilbert, Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen, Erste Mitteilung, Göttingen Nachrichten, (1904) 49–91.
- 1904b David Hilbert, Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen, Zweite Mitteilung, Göttingen Nachrichten, (1904) 213–259.
- 1904c Henri Lebesgue, Leçons sur l'intégration et la recherche des fonctions primitives, Gauthier-Villars, Paris, 1904.
- 1905 David Hilbert, Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen, Dritte Mitteilung, Göttingen Nachrichten, (1905) 307–338.
- 1906a David Hilbert, Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen, Vierte Mitteilung, Göttingen Nachrichten, (1906) 157–227.
- 1906b David Hilbert, Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen, Fünfte Mitteilung, Göttingen Nachrichten, (1906) 439–480.
- 1907a Erhard Schmidt, Zur Theorie der linearen und nichtlinearen Integralgleichungen, I, Math. Annalen, 63 (1907) 433–476.
- 1907b Erhard Schmidt, Zur Theorie der linearen und nichtlinearen Integralgleichungen, II, Math. Annalen, 64 (1907) 161–174.
- 1907c Friedrich Riesz, Sur les systèmes orthogonaux de fonctions, C. R. Acad. Sci. Paris, 144 (1907) 615–619.
- 1907d Friedrich Riesz, Über orthogonale Funktionensysteme, Göttingen Nachrichten, (1907) 116–122.
- 1907e Ernst Fischer, Sur la convergence en moyenne, C. R. Acad. Sci. Paris, 144 (1907) 1022–1024.
- 1908a Erhard Schmidt, Über die Auflösung linearer Gleichungen mit unendlich vielen Unbekannten, Rend. Circ. Mat. Palermo, 25 (1908) 53–77.
- 1908b Eliakim Hastings Moore, On a form of general analysis with application to linear differential and integral equations, Cong. Int. d. Math., Rome, 1908, 98–114.
- 1908c George D. Birkhoff, Boundary value and expansion problems of ordinary linear differential equations, Trans. Amer. Math. Soc., 9 (1908) 373–395, (Coll. Papers, I, 14–36).
- 1909a Ernst Hellinger, Neue Begründung der Theorie quadratischer Formen von unendlichvielen Veränderlichen, J. Reine Angew. Math., 136 (1909) 210–271.
- 1909b Friedrich Riesz, Sur les opérations fonctionnelles linéaires, C. R. Acad. Sci. Paris, 149 (1909) 974–77.
- 1909c Isaac Schur, Über die charakteristischen Wurzeln einer linearen Substitution mit einer Anwendung auf die Theorie der Integralgleichungen, Math. Annalen, 66 (1909) 488–510.
- 1909d Hermann Weyl, Über beschränkte quadratische Formen deren Differenz vollstetig ist, Rend. Circ. Mat. Palermo, 27 (1909) 373–392.
- 1910a David Hilbert, Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen, Sechste Mitteilungen, Göttingen Nachrichten, (1910) 355–419.
- 1910b Ernst Hellinger and Otto Toeplitz, Grundlagen für eine Theorie der unendlichen Matrizen, Math. Annalen, 69 (1910) 289–330.
- 1910c Friedrich Riesz, Untersuchungen über Systeme integrierbarer Funktionen, Math. Annalen, 69 (1910) 449–497.
- 1910d Helge von Koch, Sur les systèmes d'une infinité d'équations linéaires à une infinité d'inconnues, C. R. Cong. d. Math., Stockholm, 1910, 43–61.
- 1910e Ivar Fredholm, Les équations intégrales linéaires, C. R. Cong. d. Math. Stockholm, 1910, 92–100.
- 1910f Henri Lebesgue, Sur l'intégrale de Stieltjes et sur les opérations linéaires, C. R. Acad. Sci. Paris, 150 (1910) 86–88.
- 1912a David Hilbert, Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen, Teubner, Leipzig, 1912.

- 1912b Eliakim Hastings Moore, On the foundations of the theory of linear integral equations, *Bull. Amer. Math. Soc.*, 18 (1912) 334–362.
- 1913a Frédéric Riesz, Les systèmes d'équations linéaires à une infinité d'inconnues, Gauthier Villars, Paris, 1913.
- 1913b Johann Radon, Theorie und Anwendungen der absolut additiven Mengenfunktionen, *Sitz. Akad. Wiss., Wien*, 122 (1913) 1295–1438.
- 1918a Otto Toeplitz, Das algebraische Analogon zu einem Satze von Fejér, *Math. Zeit.*, 2 (1918) 187–197.
- 1918b Friedrich Riesz, Über lineare Funktionalgleichungen, *Acta Mathematica*, 41 (1918) 71–98.
- 1923 Torsten Carleman, Sur les équations intégrales singulières à noyau réel et symétrique, *Uppsala Univ. Årsskrift*, 1923.
- 1925a Max Born and Pascual Jordan, Zur Quantenmechanik, *Z. Physik*, 34 (1925) 858–888.
- 1925b Paul A. M. Dirac, The fundamental equations of quantum mechanics, *Proc. Roy. Soc. London, Ser. A.*, 109 (1925) 642–658.
- 1926a Max Born, Werner Heisenberg and Pascual Jordan, Zur Quantenmechanik, II, *Z. Physik*, 35 (1926) 557–615.
- 1926b Paul A. M. Dirac, On the theory of quantum mechanics, *Proc. Roy. Soc. London, Ser. A.*, 112 (1926) 661–677.
- 1927 John von Neumann, Mathematische Begründung der Quantenmechanik, *Göttingen Nachrichten* (1927) 1–57, (Coll. Works, I, 151–207).
- 1928a Ernst Hellinger and Otto Toeplitz, Integralgleichungen und Gleichungen mit unendlichvielen Unbekannten, Teubner, Leipzig, 1928.
- 1928b Norbert Wiener, The spectrum of an arbitrary function, *Proc. London Math. Soc.*, 27 (1928) 483–496.
- 1929a John von Neumann, Allgemeine Eigenwerttheorie Hermitescher Funktionaloperation, *Math. Annalen*, 102 (1929) 49–131, (Coll. Works, II, 3–85).
- 1929b Aurel Wintner, Spektraltheorie der unendlichen Matrizen, Hirzel, Leipzig, 1929.
- 1930a John von Neumann, Zur Algebra der Funktionaloperationen und Theorie der normalen Operatoren, *Math. Annalen*, 102 (1930) 370–427, (Coll. Works, II, 86–143).
- 1930b Paul A. M. Dirac, *The Principles of Quantum Mechanics*, Oxford, 1930.
- 1930c Frederich Riesz, Über die linearen Transformationen des komplexen Hilbertschen Raumes, *Acta Litt. Sci. (Szeged)*, 5 (1930) 23–54.
- 1930b Norbert Wiener, Generalized harmonic analysis, *Acta Mathematica*, 55 (1930) 117–258.
- 1931 T. H. Hildebrandt, Linear functional transformations in general spaces, *Bull. Amer. Math. Soc.*, 37 (1931) 185–212.
- 1932a Marshall H. Stone, Linear Transformations in Hilbert Space, *Amer. Math. Soc. Colloq. Publ.* XV, New York, 1932.
- 1932b Stefan Banach, *Théorie des opérations linéaires*, Warsaw, 1932.
- 1934 Kurt O. Friedrichs, Spektraltheorie halbbeschränkter Operatoren und Anwendung auf die Spektralzerlegung von Differentialoperatoren, *Math. Annalen*, 109 (1934) 465–487, 685–713.
- 1936a Francis J. Murray and John von Neumann, On rings of operators, I, *Ann. Math.*, 37 (1936) 116–229, (Von Neumann Coll. Works III, 6–119).
- 1936b S. W. P. Steen, An introduction to the theory of operators, I, *Proc. London Math. Soc.*, 41 (1936) 361–392.
- 1936c Mitio Nagumo, Einige analytische Untersuchungen in linearen metrischen Ringen, *Jap. J. Math.*, 13(1936) 61–80.
- 1937a Francis J. Murray and John von Neumann, On rings of operators, II, *Trans. Amer. Math. Soc.*, 41 (1937) 208–248 (Von Neumann Coll. Works, III, 120–160).

- 1937b S. W. P. Steen, An introduction to the theory of operators, II, *Proc. London Math. Soc.*, 43 (1937) 529–543.
- 1938a S. W. P. Steen, An introduction to the theory of operators, III, *Proc. London Math. Soc.*, 44 (1938) 398–441.
- 1938b Stanislaw Mazur, Sur les anneaux linéaires, *C. R. Acad. Sci. Paris*, 207 (1938) 1025–1027.
- 1939 S. W. P. Steen, An introduction to the theory of operators, IV, *Proc. Camb. Phil. Soc.*, 35(1939) 562–578.
- 1940a John von Neumann, On rings of operators, III, *Ann. Math.*, 41 (1940) 94–161, (Coll. Works, III, 161–228).
- 1940b S. W. P. Steen, An introduction to the theory of operators, V, *Proc. Camb. Phil. Soc.*, 36(1940) 139–149.
- 1940c Marshall H. Stone, A general theory of spectra I, *Proc. Nat. Acad. Sci. U. S. A.*, 26 (1940) 280–283.
- 1941a Israel M. Gelfand, Normierte Ringe, *Mat. Sbornik, N. S.*, 9 (1941) 3–24.
- 1941b Israel M. Gelfand and Georgii E. Silov, Über verschiedene Methoden der Einführung der Topologie in die Menge der Maximalen Ideale eines normierten Ringes, *Mat. Sbornik, N. S.*, 9 (1941) 25–38.
- 1941c Marshall H. Stone, A general theory of spectra II, *Proc. Nat. Acad. Sci. U. S. A.*, 27 (1941) 83–87.
- 1941d Shi uo Kakutani, Concrete representation of abstract (L) spaces and the mean ergodic theorem, *Ann. Math.*, 42 (1941) 523–537.
- 1942 Edgar R. Lorch, The spectrum of linear transformations, *Trans. Amer. Math. Soc.*, 52 (1942) 238–248.
- 1943 Israel M. Gelfand and M. A. Naimark, On the imbedding of normed rings into the ring of operators in Hilbert space, *Mat. Sbornik, N. S.*, 12 (1943) 197–213.
- 1944 Hermann Weyl, David Hilbert and his mathematical work, *Bull. Amer. Math. Soc.*, 50 (1944) 612–654.
- 1946 Charles E. Rickart, Banach algebras with an adjoint operation, *Ann. Math.*, 47 (1946) 528–550.
- 1949a John von Neumann, On Rings of Operators, Reduction Theory, *Ann. Math.*, 50 (1949) 401–485, (Coll. Works, III, 400–484).
- 1949b Arne Beurling, On the spectral synthesis of bounded functions, *Acta Mathematica*, 81 (1949) 225–238.
- 1951 Paul R. Halmos, *Introduction to Hilbert Space*, Chelsea, New York, 1951.
- 1953 Richard G. Cooke, *Linear Operators*, Macmillan, London, 1953.
- 1958a Nelson Dunford and Jacob T. Schwartz, *Linear Operators I: General Theory*, Interscience, New York, 1958.
- 1958b Paul R. Halmos, *Finite Dimensional Vector Spaces*, Second Edition, Van Nostrand, Princeton, N. J., 1958.
- 1962 Edgar R. Lorch, *Spectral Theory*, Oxford U. P., New York, 1962.
- 1963 Nelson Dunford and Jacob T. Schwartz, *Linear Operators II: Spectral Theory*, Interscience, New York, 1963.
- 1966 Michael Bernkopf, The Development of Function Spaces with Particular Reference to their Origins in the Integral Equation Theory, *Arch. Hist. Exact. Sci.*, 3 (1966) 1–96.
- 1968 Michael Bernkopf, A History of Infinite Matrices, *Arch. Hist. Exact Sci.*, 4 (1968) 308–358.
- 1970 Constance Reid, *Hilbert*, Springer-Verlag, Berlin, 1970.
- 1971 Nelson Dunford and Jacob T. Schwartz, *Linear Operators III: Spectral Operators*, Wiley, New York, 1971.

THE LEGEND OF JOHN VON NEUMANN

P. R. HALMOS, Indiana University

John von Neumann was a brilliant mathematician who made important contributions to quantum physics, to logic, to meteorology, to war, to the theory and applications of high-speed computing machines, and, via the mathematical theory of games of strategy, to economics.

Youth. He was born December 28, 1903, in Budapest, Hungary. He was the eldest of three sons in a well-to-do Jewish family. His father was a banker who received a minor title of nobility from the Emperor Franz Josef; since the title was hereditary, von Neumann's full Hungarian name was Margittai Neumann János. (Hungarians put the family name first. Literally, but in reverse order, the name means John Neumann of Margitta. The "of", indicated by the final "i", is where the "von" comes from; the place name was dropped in the German translation. In ordinary social intercourse such titles were never used, and by the end of the first world war their use had gone out of fashion altogether. In Hungary von Neumann is and always was known as Neumann János and his works are alphabetized under N. Incidentally, his two brothers, when they settled in the U.S., solved the name problem differently. One of them reserves the title of nobility for ceremonial occasions only, but, in daily life, calls himself Neumann; the other makes it less conspicuous by amalgamating it with the family name and signs himself Vonneuman.)

Even in the city and in the time that produced Szilárd (1898), Wigner (1902), and Teller (1908), von Neumann's brilliance stood out, and the legends about him started accumulating in his childhood. Many of the legends tell about his memory. His love of history began early, and, since he remembered what he learned, he ultimately became an expert on Byzantine history, the details of the trial of Joan of Arc, and minute features of the battles of the American Civil War.

Paul Halmos claims that he took up mathematics because he flunked his master's orals in philosophy.

He received his Univ. of Illinois Ph.D. under J.L. Doob. Then he was von Neumann's assistant, followed by positions at Illinois, Syracuse, M. I. T.'s Radiation Lab, Chicago, Michigan, Hawaii, and now is Distinguished Professor at Indiana Univ. He spent leaves at the Univ. of Uruguay, Montevideo, Univ. of Miami, Univ. of California, Berkeley, Tulane, and Univ. of Washington. He held a Guggenheim Fellowship and was awarded the MAA Chauvenet Prize.

Professor Halmos' research is mainly measure theory, probability, ergodic theory, topological groups, Boolean algebra, algebraic logic, and operator theory in Hilbert space. He has served on the Council of the AMS for many years and was Editor of the Proceedings of the AMS and Mathematical Reviews. His eight books, all widely used, include *Finite-Dimensional Vector Spaces* (Van Nostrand, 1958), *Measure Theory* (Van Nostrand, 1950), *Naive Set Theory* (Van Nostrand, 1960), and *Hilbert Space Problem Book* (Van Nostrand, 1967).

The present paper is the original uncut version of a brief article commissioned by the Encyclopaedia Britannica. *Editor.*

He could, it is said, memorize the names, addresses, and telephone numbers in a column of the telephone book on sight. Some of the later legends tell about his wit and his fondness for humor, including puns and off-color limericks. Speaking of the Manhattan telephone book he said once that he knew all the numbers in it — the only other thing he needed, to be able to dispense with the book altogether, was to know the names that the numbers belonged to.

Most of the legends, from childhood on, tell about his phenomenal speed in absorbing ideas and solving problems. At the age of 6 he could divide two eight-digit numbers in his head; by 8 he had mastered the calculus; by 12 he had read and understood Borel's *Théorie des Fonctions*.

These are some of the von Neumann stories in circulation. I'll report others, but I feel sure that I haven't heard them all. Many are undocumented and unverifiable, but I'll not insert a separate caveat for each one: let this do for them all. Even the purely fictional ones say something about him; the stories that men make up about a folk hero are, at the very least, a strong hint to what he was like.)

In his early teens he had the guidance of an intelligent and dedicated high-school teacher, L. Rátz, and, not much later, he became a pupil of the young M. Fekete and the great L. Fejér, "the spiritual father of many Hungarian mathematicians". ("Fekete" means "Black", and "Fejér" is an archaic spelling, analogous to "Whyte".)

According to von Kármán, von Neumann's father asked him, when John von Neumann was 17, to dissuade the boy from becoming a mathematician, for financial reasons. As a compromise between father and son, the solution von Kármán proposed was chemistry. The compromise was adopted, and von Neumann studied chemistry in Berlin (1921–1923) and in Zürich (1923–1925). In 1926 he got both a Zürich diploma in chemical engineering and a Budapest Ph.D. in mathematics.

Early work. His definition of ordinal numbers (published when he was 20) is the one that is now universally adopted. His Ph.D. dissertation was about set theory too; his axiomatization has left a permanent mark on the subject. He kept up his interest in set theory and logic most of his life, even though he was shaken by K. Gödel's proof of the impossibility of proving that mathematics is consistent.

He admired Gödel and praised him in strong terms: "Kurt Gödel's achievement in modern logic is singular and monumental — indeed it is more than a monument, it is a landmark which will remain visible far in space and time. ... The subject of logic has certainly completely changed its nature and possibilities with Gödel's achievement." In a talk entitled "The Mathematician", speaking, among other things, of Gödel's work, he said: "This happened in our lifetime, and I know myself how humiliatingly easily my own values regarding the absolute mathematical truth changed during this episode, and how they changed three times in succession!"

He was Privatdozent at Berlin (1926–1929) and at Hamburg (1929–1930). During this time he worked mainly on two subjects, far from set theory but near to one another: quantum physics and operator theory. It is almost not fair to call them two

subjects: due in great part to von Neumann's own work, they can be viewed as two aspects of the same subject. He started the process of making precise mathematics out of quantum theory, and (it comes to the same thing really) he was inspired by the new physical concepts to make broader and deeper the purely mathematical study of infinite-dimensional spaces and operators on them. The basic insight was that the geometry of the vectors in a Hilbert space has the same formal properties as the structure of the states of a quantum-mechanical system. Once that is accepted, the difference between a quantum physicist and a mathematical operator-theorist becomes one of language and emphasis only. Von Neumann's book on quantum mechanics appeared (in German) in 1932. It has been translated into French (1947), Spanish (1949), and English (1955), and it is still one of the standard and one of the most inspiring treatments of the subject. Speaking of von Neumann's contributions to quantum mechanics, E. Wigner, a Nobel laureate, said that they alone "would have secured him a distinguished position in present day theoretical physics".

Princeton. In 1930 von Neumann went to Princeton University for one term as visiting lecturer, and the following year he became professor there. In 1933, when the Institute for Advanced Study was founded, he was one of the original six professors of its School of Mathematics, and he kept that position for the rest of his life. (It is easy to get confused about the Institute and its formal relation with Princeton University, even though there is none. They are completely distinct institutions. The Institute was founded for scholarship and research only, not teaching. The first six professors in the School of Mathematics were J. W. Alexander, A. Einstein, M. Morse, O. Veblen, J. von Neumann, and H. Weyl. When the Institute began it had no building, and it accepted the hospitality of Princeton University. Its members and visitors have, over the years, maintained close professional and personal relations with their colleagues at the University. These facts kept contributing to the confusion, which was partly clarified in 1940, when the Institute acquired a building of its own, about a mile from the Princeton campus.)

In 1930 von Neumann married Marietta Kövesi; in 1935 their daughter Marina was born. (In 1956 Marina von Neumann graduated from Radcliffe *summa cum laude*, with the highest scholastic record in her class. In 1972 Marina von Neumann Whitman was appointed by President Nixon to the Council of Economic Advisers.) In the 1930's the stature of von Neumann, the mathematician, grew at the rate that his meteoric early rise had promised, and the legends about Johnny, the human being, grew along with it. He enjoyed life in America and lived it in an informal manner, very differently from the style of the conventional German professor. He was not a refugee and he didn't feel like one. He was a cosmopolite in attitude and a U.S. citizen by choice.

The parties at the von Neumanns' house were frequent, and famous, and long. Johnny was not a heavy drinker, but he was far from a teetotaler. In a roadside

restaurant he once ordered a brandy with a hamburger chaser. The outing was in honor of his birthday and he was feeling fine that evening. One of his gifts was a toy, a short prepared tape attached to a cardboard box that acted as sounding board; when the tape was pulled briskly past a thumbnail, it would squawk "Happy birthday!" Johnny squawked it often. Another time, at a party at his house, there was one of those thermodynamic birds that dips his beak in a glass of water, straightens up, teeter-totters for a while, and then repeats the cycle. A temporary but firm house rule was quickly passed: everyone had to take a drink each time that the bird did.

He liked to drive, but he didn't do it well. There was a "von Neumann's corner" in Princeton, where, the story goes, his cars repeatedly had trouble. One often quoted explanation that he allegedly offered for one particular crack-up goes like this: "I was proceeding down the road. The trees on the right were passing me in orderly fashion at 60 miles an hour. Suddenly one of them stepped in my path. Boom!"

He once had a dog named "Inverse". He played poker, but only rarely, and he usually lost.

In 1937 the von Neumanns were divorced; in 1938 he married Klára Dán. She learned mathematics from him and became an expert programmer. Many years later, in an interview, she spoke about him. "He has a very weak idea of the geography of the house. ...Once, in Princeton, I sent him to get me a glass of water; he came back after a while wanting to know where the glasses were. We had been in the house only seventeen years. ...He has never touched a hammer or a screwdriver; he does nothing around the house. Except for fixing zippers. He can fix a broken zipper with a touch."

Von Neumann was definitely not the caricatured college professor. He was a round, pudgy man, always neatly, formally dressed. There are, to be sure, one or two stories of his absentmindedness. Klári told one about the time when he left their Princeton house one morning to drive to a New York appointment, and then phoned her when he reached New Brunswick to ask: "Why am I going to New York?" It may not be strictly relevant, but I am reminded of the time I drove him to his house one afternoon. Since there was to be a party there later that night, and since I didn't trust myself to remember exactly how I got there, I asked how I'd be able to know his house when I came again. "That's easy," he said; "it's the one with that pigeon sitting by the curb."

Normally he was alert, good at rapid repartee. He could be blunt, but never stuffy, never pompous. Once the telephone interrupted us when we were working in his office. His end of the conversation was very short; all he said between "Hello" and "Goodbye" was "Fekete pestis!", which means "Black plague!" Remembering, after he hung up, that I understood Hungarian, he turned to me, half apologetic and half exasperated, and explained that he wasn't speaking of one of the horsemen of the Apocalypse, but merely of some unexpected and unwanted dinner guests that his wife just told him about.

On a train once, hungry, he asked the conductor to send the man with the sandwich

tray to his seat. The busy and impatient conductor said "I will if I see him". Johnny's reply: "This train is linear, isn't it?"

Speed. The speed with which von Neumann could think was awe-inspiring. G. Pólya admitted that "Johnny was the only student I was ever afraid of. If in the course of a lecture I stated an unsolved problem, the chances were he'd come to me as soon as the lecture was over, with the complete solution in a few scribbles on a slip of paper." Abstract proofs or numerical calculations — he was equally quick with both, but he was especially pleased with and proud of his facility with numbers. When his electronic computer was ready for its first preliminary test, someone suggested a relatively simple problem involving powers of 2. (It was something of this kind: what is the smallest power of 2 with the property that its decimal digit fourth from the right is 7? This is a completely trivial problem for a present-day computer: it takes only a fraction of a second of machine time.) The machine and Johnny started at the same time, and Johnny finished first.

One famous story concerns a complicated expression that a young scientist at the Aberdeen Proving Ground needed to evaluate. He spent ten minutes on the first special case; the second computation took an hour of paper and pencil work; for the third he had to resort to a desk calculator, and even so took half a day. When Johnny came to town, the young man showed him the formula and asked him what to do. Johnny was glad to tackle it. "Let's see what happens for the first few cases. If we put $n = 1$, we get..." — and he looked into space and mumbled for a minute. Knowing the answer, the young questioner put in " 2.31 ?" Johnny gave him a funny look and said "Now if $n = 2$,..." and once again voiced some of his thoughts as he worked. The young man, prepared, could of course follow what Johnny was doing, and, a few seconds before Johnny finished, he interrupted again, in a hesitant tone of voice: " 7.49 ?" This time Johnny frowned, and hurried on: "If $n = 3$, then...". The same thing happened as before — Johnny muttered for several minutes, the young man eavesdropped, and, just before Johnny finished, the young man exclaimed: " 11.06 !" That was too much for Johnny. It couldn't be! No unknown beginner could outdo him! He was upset and he sulked till the practical joker confessed.

Then there is the famous fly puzzle. Two bicyclists start twenty miles apart and head toward each other, each going at a steady rate of 10 m.p.h. At the same time a fly that travels at a steady 15 m.p.h. starts from the front wheel of the southbound bicycle and flies to the front wheel of the northbound one, then turns around and flies to the front wheel of the southbound one again, and continues in this manner till he is crushed between the two front wheels. Question: what total distance did the fly cover? The slow way to find the answer is to calculate what distance the fly covers on the first, northbound, leg of the trip, then on the second, southbound, leg, then on the third, etc., etc., and, finally, to sum the infinite series so obtained. The quick way is to observe that the bicycles meet exactly one hour after their start, so that the fly had just an hour for his travels; the answer must therefore be 15 miles. When the

question was put to von Neumann, he solved it in an instant, and thereby disappointed the questioner: "Oh, you must have heard the trick before!" "What trick?" asked von Neumann; "all I did was sum the infinite series."

I remember one lecture in which von Neumann was talking about rings of operators. At an appropriate point he mentioned that they can be classified two ways: finite versus infinite, and discrete versus continuous. He went on to say: "This leads to a total of four possibilities, and, indeed, all four of them can occur. Or — let's see — can they?" Many of us in the audience had been learning this subject from him for some time, and it was no trouble to stop and mentally check off all four possibilities. No trouble — it took something like two seconds for each, and, allowing for some fumbling and shifting of gears, it took us perhaps 10 seconds in all. But after two seconds von Neumann had already said "Yes, they can," and he was two sentences into the next paragraph before, dazed, we could scramble aboard again.

Speech. Since Hungarian is not exactly a *lingua franca*, all educated Hungarians must acquire one or more languages with a popular appeal greater than that of their mother tongue. At home the von Neumanns spoke Hungarian, but he was perfectly at ease in German, and in French, and, of course, in English. His English was fast and grammatically defensible, but in both pronunciation and sentence construction it was reminiscent of German. His "Sprachgefühl" was not perfect, and his sentences tended to become involved. His choice of words was usually exactly right; the occasional oddities (like "a self-obvious theorem") disappeared in later years. His spelling was sometimes more consistent than commonplace: if "commit", then "ommit". S. Ulam tells about von Neumann's trip to Mexico, where "he tried to make himself understood by using 'neo-Castilian', a creation of his own — English words with an 'el' prefix and appropriate Spanish endings".

He prepared for lectures, but rarely used notes. Once, five minutes before a non-mathematical lecture to a general audience, I saw him as he was preparing. He sat in the lounge of the Institute and scribbled on a small card a few phrases such as these: "Motivation, 5 min.; historical background, 15 min.; connection with economics, 10 min.;..."

As a mathematical lecturer he was dazzling. He spoke rapidly but clearly; he spoke precisely, and he covered the ground completely. If, for instance, a subject has four possible axiomatic approaches, most teachers content themselves with developing one, or at most two, and merely mentioning the others. Von Neumann was fond of presenting the "complete graph" of the situation. He would, that is, describe the shortest path that leads from the first to the second, from the first to the third, and so on through all twelve possibilities.

His one irritating lecturing habit was the way he wielded an eraser. He would write on the board the crucial formula under discussion. When one of the symbols in it had been proved to be replaceable by something else, he made the replacement not by rewriting the whole formula, suitably modified, but by erasing the replaceable

symbol and substituting the new one for it. This had the tendency of inducing symptoms of acute discouragement among note-takers, especially since, to maintain the flow of the argument, he would keep talking at the same time.

His style was so persuasive that one didn't have to be an expert to enjoy his lectures; everything seemed easy and natural. Afterward, however, the Chinese-dinner phenomenon was likely to occur. A couple of hours later the average memory could no longer support the delicate balance of mutually interlocking implications, and, puzzled, would feel hungry for more explanation.

Style. As a writer of mathematics von Neumann was clear, but not clean; he was powerful but not elegant. He seemed to love fussy detail, needless repetition, and notation so explicit as to be confusing. To maintain a logically valid but perfectly transparent and unimportant distinction, in one paper he introduced an extension of the usual functional notation: along with the standard $\phi(x)$ he dealt also with something denoted by $\phi((x))$. The hair that was split to get there had to be split again a little later, and there was $\phi(((x)))$, and, ultimately, $\phi((((x))))$. Equations such as

$$(\psi((((a))))))^2 = \phi((((a))))$$

have to be peeled before they can be digested; some irreverent students referred to this paper as von Neumann's onion.

Perhaps one reason for von Neumann's attention to detail was that he found it quicker to hack through the underbrush himself than to trace references and see what others had done. The result was that sometimes he appeared ignorant of the standard literature. If he needed facts, well-known facts, from Lebesgue integration theory, he waded in, defined the basic notions, and developed the theory to the point where he could use it. If, in a later paper, he needed integration theory again, he would go back to the beginning and do the same thing again.

He saw nothing wrong with long strings of suffixes, and subscripts on subscripts; his papers abound in avoidable algebraic computations. The reason, probably, is that he saw the large picture; the trees did not conceal the forest from him. He saw and he relished all parts of the mathematics he was thinking about. He never wrote "down" to an audience; he told it as he saw it. The practice caused no harm; the main result was that, quite a few times, it gave lesser men an opportunity to publish "improvements" of von Neumann.

Since he had no formal connections with educational institutions after he was 30, von Neumann does not have a long list of students; he supervised only one Ph.D. thesis. Through his lectures and informal conversations he acquired, however, quite a few disciples who followed in one or another of his footsteps. A few among them are J. W. Calkin, J. Charney, H. H. Goldstine, P. R. Halmos, I. Halperin, O. Morgenstern, F. J. Murray, R. Schatten, I. E. Segal, A. H. Taub, and S. Ulam.

Work habits. Von Neumann was not satisfied with seeing things quickly and clearly; he also worked very hard. His wife said "he had always done his writing at home during the night or at dawn. His capacity for work was practically unlimited." In addition to his work at home, he worked hard at his office. He arrived early, he stayed late, and he never wasted any time. He was systematic in both large things and small; he was, for instance, a meticulous proofreader. He would correct a manuscript, record on the first page the page numbers where he found errors, and, by appropriate tallies, record the number of errors that he had marked on each of those pages. Another example: when requested to prepare an abstract of not more than 200 words, he would not be satisfied with a statistical check — there are roughly 20 lines with about 10 words each — but he would count every word.

When I was his assistant we wrote one paper jointly. After the thinking and the talking were finished, it became my job to do the writing. I did it, and I submitted to him a typescript of about 12 pages. He read it, criticized it mercilessly, crossed out half, and rewrote the rest; the result was about 18 pages. I removed some of the Germanisms, changed a few spellings, and compressed it into 16 pages. He was far from satisfied, and made basic changes again; the result was 20 pages. The almost divergent process continued (four innings on each side as I now recall it); the final outcome was about 30 typescript pages (which came to 19 in print).

Another notable and enviable trait of von Neumann's was his mathematical courage. If, in the middle of a search for a counterexample, an infinite series came up, with a lot of exponentials that had quadratic exponents, many mathematicians would start with a clean sheet of paper and look for another counterexample. Not Johnny! When that happened to him, he cheerfully said: "Oh, yes, a theta function...", and plowed ahead with the mountainous computations. He wasn't afraid of anything.

He knew a lot of mathematics, but there were also gaps in his knowledge, most notably number theory and algebraic topology. Once when he saw some of us at a blackboard staring at a rectangle that had arrows marked on each of its sides, he wanted to know that what was. "Oh just the torus, you know — the usual identification convention." No, he didn't know. The subject is elementary, but some of it just never crossed his path, and even though most graduate students knew about it, he didn't.

Brains, speed, and hard work produced results. In von Neumann's *Collected Works* there is a list of over 150 papers. About 60 of them are on pure mathematics (set theory, logic, topological groups, measure theory, ergodic theory, operator theory, and continuous geometry), about 20 on physics, about 60 on applied mathematics (including statistics, game theory, and computer theory), and a small handful on some special mathematical subjects and general non-mathematical ones. A special number of the *Bulletin of the American Mathematical Society* was devoted to a discussion of his life and work (in May 1958).

Pure mathematics. Von Neumann's reputation as a mathematician was firmly

established by the 1930's, based mainly on his work on set theory, quantum theory, and operator theory, but enough more for about three ordinary careers, in pure mathematics alone, was still to come. The first of these was the proof of the ergodic theorem. Various more or less precise statements had been formulated earlier in statistical mechanics and called the ergodic hypothesis. In 1931 B. O. Koopman published a penetrating remark whose main substance was that one of the contexts in which a precise statement of the ergodic hypothesis could be formulated is the theory of operators on Hilbert space — the very subject that von Neumann used earlier to make quantum mechanics precise and on which he had written several epoch-making papers. It is tempting to speculate on von Neumann's reaction to Koopman's paper. It could have been something like this: "By Koopman's remark the ergodic hypothesis becomes a theorem about Hilbert spaces — and if that's what it is I ought to be able to prove it. Let's see now..." Soon after the appearance of Koopman's paper, von Neumann formulated and proved the statement that is now known as the mean ergodic theorem for unitary operators. There was some temporary confusion, caused by publication dates, about who did what before whom, but by now it is universally recognized that von Neumann's theorem preceded and inspired G. D. Birkhoff's point ergodic theorem. In the course of the next few years von Neumann published several more first-rate papers on ergodic theory, and he made use of the techniques and results of that theory later, in his studies of rings of operators.

In 1900 D. Hilbert presented a famous list of 23 problems that summarized the state of mathematical knowledge at the time and showed where further work was needed. In 1933 A. Haar proved the existence of a suitable measure (which has come to be called Haar measure) in topological groups; his proof appears in the *Annals of Mathematics*. Von Neumann had access to Haar's result before it was published, and he quickly saw that that was exactly what was needed to solve an important special case (compact groups) of one of Hilbert's problems (the 5th). His solution appears in the same issue of the same journal, immediately after Haar's paper.

In the second half of the 1930's the main part of von Neumann's publications was a sequence of papers, partly in collaboration with F. J. Murray, on what he called rings of operators. (They are now called von Neumann algebras.) It is possible that this is the work for which von Neumann will be remembered the longest. It is a technically brilliant development of operator theory that makes contact with von Neumann's earlier work, generalizes many familiar facts about finite-dimensional algebra, and is currently one of the most powerful tools in the study of quantum physics.

A surprising outgrowth of the theory of rings of operators is what von Neumann called continuous geometry. Ordinary geometry deals with spaces of dimension 1, 2, 3, etc. In his work on rings of operators von Neumann saw that what really determines the dimension structure of a space is the group of rotations that it admits. The group of rotations associated with the ring of *all* operators yields the familiar dimensions. Other groups, associated with different rings, assign to spaces dimensions

whose values can vary continuously; in that context it makes sense to speak of a space of dimension $3/4$, say. Abstracting from the “concrete” case of rings of operators, von Neumann formulated the axioms that make these continuous-dimensional spaces possible. For several years he thought, wrote, and lectured about continuous geometries. In 1937 he was the Colloquium Lecturer of the American Mathematical Society and chose that subject for his topic.

Applied mathematics. The year 1940 was just about the half-way point of von Neumann’s scientific life, and his publications show a discontinuous break then. Till then he was a topflight pure mathematician who understood physics; after that he was an applied mathematician who remembered his pure work. He became interested in partial differential equations, the principal classical tool of the applications of mathematics to the physical world. Whether the war made him into an applied mathematician or his interest in applied mathematics made him invaluable to the war effort, in either case he was much in demand as a consultant and advisor to the armed forces and to the civilian agencies concerned with the problems of war. His papers from this point on are mainly on statistics, shock waves, flow problems, hydrodynamics, aerodynamics, ballistics, problems of detonation, meteorology, and, last but not least, two non-classical, new aspects of the applicability of mathematics to the real world: games and computers.

Von Neumann’s contributions to war were manifold. Most often mentioned is his proposal of the implosion method for bringing nuclear fuel to explosion (during World War II) and his espousal of the development of the hydrogen bomb (after the war). The citation that accompanied his honorary D.Sc. from Princeton in 1947 mentions (in one word) that he was a mathematician, but praises him for being a physicist, an engineer, an armorer, and a patriot.

Politics. His political and administrative decisions were rarely on the side that is described nowadays by the catchall term “liberal”. He appeared at times to advocate preventive war with Russia. As early as 1946 atomic bomb tests were already receiving adverse criticism, but von Neumann thought that they were necessary and (in, for instance, a letter to the *New York Times*) defended them vigorously. He disagreed with J. R. Oppenheimer on the H-bomb crash program, and urged that the U.S. proceed with it before Russia could. He was, however, a “pro-Oppenheimer” witness at the Oppenheimer security hearings. He said that Oppenheimer opposed the program “in good faith” and was “very constructive” once the decision to go ahead with the super bomb was made. He insisted that Oppenheimer was loyal and was not a security risk.

As a member of the Atomic Energy Commission (appointed by President Eisenhower, he was sworn in on March 15, 1955), having to “think about the unthinkable”, he urged a United Nations study of world-wide radiation effects. “We willingly pay 30,000–40,000 fatalities per year (2% of the total death rate),” he wrote, “for the advantages of individual transportation by automobile.” He mentioned a

fall-out accident in an early Pacific bomb test that resulted in one fatality and danger to 200 people, and he compared it with a Japanese ferry accident that “killed about 1,000 people, including 20 Americans — yet the...fall-out was what attracted almost world-wide attention.” He asked: “Is the price in international popularity worth paying?” And he answered: “Yes: we have to accept it as part payment for our more advanced industrial position.”

Game theory. At about the same time that he began to apply his analytic talents to the problems of war, von Neumann found time and energy to apply his combinatorial insight to what he called the theory of games, whose major application was to economics. The mathematical cornerstone of the theory is one statement, the so-called minimax theorem, that von Neumann proved early (1928) in a short article (25 pages); its elaboration and applications are in the book he wrote jointly with O. Morgenstern in 1944. The minimax theorem says about a large class of two-person games that there is no point in playing them. If either player considers, for each possible strategy of play, the *maximum* loss that he can expect to sustain with that strategy, and then chooses the “optimal” strategy that *minimizes* the maximum loss, then he can be statistically sure of not losing more than that minimax value. Since (and this is the whole point of the theorem) that value is the negative of the one, similarly defined, that his opponent can guarantee for himself, the long-run outcome is completely determined by the rules.

Mathematical economics before von Neumann tried to achieve success by imitating the technique of classical mathematical physics. The mathematical tools used were those of analysis (specifically the calculus of variations), and the procedure relied on a not completely reliable analogy between economics and mechanics. The secret of the success of the von Neumann approach was the abandonment of the mechanical analogy and its replacement by a fresh point of view (games of strategy) and new tools (the ideas of combinatorics and convexity).

The role that game theory will play in the future of mathematics and economics is not easy to predict. As far as mathematics is concerned, it is tenable that the only thing that makes the Morgenstern-von Neumann book 600 pages longer than the original von Neumann paper is the development needed to apply the abstruse deductions of one subject to the concrete details of another. On the other hand, enthusiastic proponents of game theory can be found who go so far as to say that it may be “one of the major scientific contributions of the first half of the 20th century”.

Machines. The last subject that contributed to von Neumann’s fame was the theory of electronic computers and automata. He was interested in them from every point of view: he wanted to understand them, design them, build them, and use them. What are the logical components of the processes that a computer will be asked to perform? What is the best way of obtaining practically reliable answers from a machine with unreliable components? What does a machine need to “remember”, and what is the best way to equip it with a “memory”? Can a machine be built that can not

only save us the labor of computing but save us also the trouble of building a new machine — is it possible, in other words, to produce a self-reproducing automaton? (Answer: in principle, yes. A sufficiently complicated machine, embedded in a thick chowder of randomly distributed spare parts, its “food”, would pick up one part after another till it found a usable one, put it in place, and continue to search and construct till its descendant was complete and operational.) Can a machine successfully imitate “randomness”, so that when no formulae are available to solve a concrete physical problem (such as that of finding an optimal bombing pattern), the machine can perform a large number of probability experiments and yield an answer that is statistically accurate? (The last question belongs to the concept that is sometimes described as the Monte Carlo method.) These are some of the problems that von Neumann studied and to whose solutions he made basic contributions.

He had close contact with several computers — among them the MANIAC (Mathematical Analyzer, Numerical Integrator, Automatic Calculator), and the affectionately named JONIAc. He advocated their use for everything from the accumulation of heuristic data for the clarification of our intuition about partial differential equations to the accurate long-range prediction and, ultimately, control of the weather. One of the most striking ideas whose study he suggested was to dye the polar icecaps so as to decrease the amount of energy they would reflect — the result could warm the earth enough to make the climate of Iceland approximate that of Hawaii.

The last academic assignment that von Neumann accepted was to deliver and prepare for publication the Silliman lectures at Yale. He worked on that job in the hospital where he died, but he couldn't finish it. His notes for it were published, and even they make illuminating reading. They contain tantalizing capsule statements of insights, and throughout them there shines an attitude of faith in and dedication to knowledge. While physicists, engineers, meteorologists, statisticians, logicians, and computers all proudly claim von Neumann as one of theirs, the Silliman lectures prove, indirectly by their approach and explicitly in the author's words, that von Neumann was first, foremost, and always a mathematician.

Death. Von Neumann was an outstanding man in tune with his times, and it is not surprising that he received many awards and honors. There is no point in listing them all here, but a few may be mentioned. He received several honorary doctorates, including ones from Princeton (1947), Harvard (1950), and Istanbul (1952). He served a term as president of the American Mathematical Society (1951–1953), and he was a member of several national scientific academies (including, of course, that of the U. S.). Somewhat to his embarrassment, he was elected to the East German Academy of Science, but the election didn't seem to take — in later years no mention is made of it in the standard biographical reference works. He received the Enrico Fermi award in 1956, when he already knew that he was incurably ill.

Von Neumann became ill in 1955. There was an operation, and the result was a diagnosis of cancer. He kept on working, and even travelling, as the disease progressed. Later he was confined to a wheelchair, but still thought, wrote, and attended meetings. In April 1956 he entered Walter Reed Hospital, and never left it. Of his last days his good friend Eugene Wigner wrote: "When von Neumann realized he was incurably ill, his logic forced him to realize that he would cease to exist, and hence cease to have thoughts. ...It was heartbreaking to watch the frustration of his mind, when all hope was gone, in its struggle with the fate which appeared to him unavoidable but unacceptable."

Von Neumann was baptized a Roman Catholic (in the U. S.), but, after his divorce, he was not a practicing member of the church. In the hospital he asked to see a priest — "one that will be intellectually compatible". Arrangements were made, he was given special instruction, and, in due course, he again received the sacraments. He died February 8, 1957.

The heroes of humanity are of two kinds: the ones who are just like all of us, but very much more so, and the ones who, apparently, have an extra-human spark. We can all run, and some of us can run the mile in less than 4 minutes; but there is nothing that most of us can do that compares with the creation of the Great G-minor Fugue. Von Neumann's greatness was the human kind. We can all think clearly, more or less, some of the time, but von Neumann's clarity of thought was orders of magnitude greater than that of most of us, all the time. Both Norbert Wiener and John von Neumann were great men, and their names will live after them, but for different reasons. Wiener saw things deeply but intuitively; von Neumann saw things clearly and logically.

What made von Neumann great? Was it the extraordinary rapidity with which he could understand and think and the unusual memory that retained everything he had once thought through? No. These qualities, however impressive they might have been, are ephemeral; they will have no more effect on the mathematics and the mathematicians of the future than the prowess of an athlete of a hundred years ago has on the sport of today.

The "axiomatic method" is sometimes mentioned as the secret of von Neumann's success. In his hands it was not pedantry but perception; he got to the root of the matter by concentrating on the basic properties (axioms) from which all else follows. The method, at the same time, revealed to him the steps to follow to get from the foundations to the applications. He knew his own strengths and he admired, perhaps envied, people who had the complementary qualities, the flashes of irrational intuition that sometimes change the direction of scientific progress. For von Neumann it seemed to be impossible to be unclear in thought or in expression. His insights were illuminating and his statements were precise.

ALTERNATING EULER PATHS FOR PACKINGS AND COVERS

C. T. ZAHN, JR., Stanford University

1. Introduction. An interesting combinatorial problem known as the "school-girls' walk" asks if the girls in an all-girl school can take a walk in two-by-two fashion so that each pair walking side by side are on friendly terms, it being known which pairs are friendly among all possible pairings. If such a utopian arrangement is not possible, then what is the largest number of friendly pairings that can be achieved simultaneously and how can such an optimal set of pairings be found? This problem is abstractly equivalent to a problem in graph theory which is as follows: Let G be a finite graph with vertex set V and edge set E ; a **matching** M of graph G is a subset of E such that no two edges in M have a vertex in common. A matching M^* is **maximum** if no other matching has more edges than M^* . A matching P is **perfect** if each vertex in V belongs to an edge of P . In this abstract version the vertices represent girls and edges represent pairs of friendly girls. A matching is a pairing of friendly girls with no girl appearing twice, an obviously necessary requirement. A perfect matching represents the utopian arrangement and a maximum matching achieves the largest number of friendly pairings.

A related problem concerns a **minimum** cover for a graph G where a **cover** C is a subset of E such that every vertex belongs to at least one edge in C . A perfect matching is then a matching which is also a cover and it is easy to see that any subset of edges which is both a matching and a cover is necessarily a maximum matching and a minimum cover.

A good algorithm for finding a maximum matching requires a reasonably simple condition which, when true, assures that a matching M is maximum and, when false, implies M is not maximum, and further indicates how to modify M to obtain a larger matching M' . Such a condition is afforded by augmenting paths. A **path** in G is a sequence of edges such that two edges adjacent in the sequence share a vertex in G . For our purposes, no edge can appear more than once in a path. If S is a subset of edges in G , an **S -alternating path** is a path whose edges are alternately in S and in $\bar{S} = E - S$. A vertex v is **S -exposed** if v belongs to no edge in S . An **M -augmenting path** for matching M is an M -alternating path whose end vertices are M -exposed. Notice this implies the end edges are in \bar{M} and the path has odd length. Such a path is called **augmenting** because by interchanging the M and \bar{M} status of the edges of the path, a new matching M' results with $|M'| = |M| + 1$. Hence, the existence

Charles Zahn studied at Princeton Univ., the Catholic Univ. of America, and the Univ. of Wisconsin. He has worked with the Applied Math Div. of NBS, the General Electric Co. — Computer Department, and presently with the Computation Group of the Stanford Linear Accelerator Center. He is a member of the MAA and the ACM, and his main research interest is picture processing and pattern recognition. *Editor.*

of an M -augmenting path implies M is not maximum. What is not so obvious is that the non-existence of an M -augmenting path implies M is maximum. This result, first obtained by Berge [1] means that the non-existence of an augmenting path is a condition which can be used to find a maximum matching. Interest in an efficient algorithm stems from several interesting practical problems which can be formulated as optimum matching or cover problems (see [2], [3], [4, p. 177] for details).

An analogous situation holds for minimum covers. A vertex v is **S -doubled** for a subset S of edges if v belongs to at least two edges of S . A **C -reducing path** for a cover C is a C -alternating path whose end edges are in C and whose end vertices are C -doubled. Once again a C -reducing path leads to a new cover C' with $|C'| = |C| - 1$. Furthermore, Norman and Rabin [2] have shown that the non-existence of a C -reducing path implies C is minimum, leading to an algorithm for finding a minimum cover. They also show that a maximum matching M^* can be obtained from a minimum cover C^* by deleting all but one C^* -edge from each C^* -doubled vertex. Adding an edge to M^* to cover each M^* -exposed vertex of a maximum matching M^* produces a minimum cover C^* . Edmonds [3] has generalized the theorems of Berge [1] and Norman-Rabin [2] by replacing edges (2-element subsets) with more general subsets of vertices whereupon the appropriate improvement structures become trees in a certain graph.

2. Packings and covers. Here we are concerned with a different generalization of the matching and cover problems. Let δ be a function which assigns a non-negative integer to each vertex v of G . If $d(v)$, the **local degree** of v , is the number of edges to which v belongs and $\delta(v) \leq d(v)$ for all v in V then δ is called a **local degree constraint** on G . A **δ -packing** in G is a subset P of edges such that each vertex v in V belongs to, at most, $\delta(v)$ edges in P . A **δ -cover** C is a subset of edges, such that each vertex v belongs to at least $\delta(v)$ edges in C . In this terminology, a matching is a 1-packing (i.e., a δ -packing with $\delta \equiv 1$) and a cover is a 1-cover. Optimum δ -packings and δ -covers are defined in the obvious way.

There is a strong duality between δ -packings and δ -covers which does not exist between matchings and covers. If δ is a local degree constraint on G , then $\bar{\delta} = d - \delta$ is also a local degree constraint on G and we say δ and $\bar{\delta}$ are **complementary**. It is not hard to see that a subset S of edges of G is a δ -packing if and only if \bar{S} is a $\bar{\delta}$ -cover. Let v be a vertex and S be a subset of edges; then v is (S, δ) -deficient; if v belongs to less than $\delta(v)$ edges in S and is (S, δ) -surfeited if v belongs to more than $\delta(v)$ edges in S . A (P, δ) -augmenting path for δ -packing P is a P -alternating path whose end edges are in \bar{P} and whose end vertices are (P, δ) -deficient. A (C, δ) -reducing path for δ -cover C is a C -alternating path whose end edges are in C and whose end vertices are (C, δ) -surfeited. In case the end vertices of an augmenting (reducing) path are identical, the deficiency (surfeit) is required to be at least two. Now the duality means that path π in G is (P, δ) -augmenting for δ -packing

P if and only if π is (\bar{P}, δ) -reducing for δ -cover \bar{P} . It is also clear that P is a maximum δ -packing if and only if $C = \bar{P}$ is a minimum δ -cover.

The theorem of Berge-Norman-Rabin proved by Berge [4, p. 175] asserts that the non-existence of a (P, δ) -augmenting path for δ -packing P implies P is maximum. Using duality, we see immediately that non-existence of a (C, δ) -reducing path for δ -cover C insures C is minimum. Goldman [5] has proved the B-N-R theorem on augmentable δ -packings by a direct reduction to the theorem of Berge [1] on augmentable matchings (1-packings). The main result of this paper is a simple direct proof of the B-N-R theorem using ideas from Edmonds' simple proof [6] of the original augmenting path theorem of Berge [1] and the notion of Euler path [7] which dates back to 1736.

As regards design of algorithms, Edmonds [6] has found an exceptionally efficient algorithm for determining a maximum matching by growing alternating-path trees and occasionally shrinking odd-length cyclic paths until an augmenting path is discovered or the edges of the graph have been depleted. Witzgall and Zahn [8] have devised a modified version of the Edmonds algorithm which does not shrink and Edmonds [9] has extended his algorithm to the case where edges have real-valued weights and maximum is defined accordingly.

The reader is referred to Berge [4] and Ore [10] for more leisurely discussions of matchings and coverings in graphs. Alternating paths were invented by Petersen [11] in the last century and augmenting paths for δ -packings occur in Tutte's [12] paper on f -factors (perfect f -packings) in a graph.

3. Euler paths. One of the earliest problems in graph theory was posed and solved by Euler [7] in the year 1736. The problem is known as the "Königsberg bridge problem" and intrigued the inhabitants of this Prussian town until solved by Euler. We quote Euler's [7] statement of the problem:

In the town of Königsberg in Prussia there is an island A, called "Kneiphof", with the two branches of the river (Pregel) flowing around it, as shown in Figure 1. There are seven bridges, **a**, **b**, **c**, **d**, **e**, **f** and **g**, crossing the two branches. The question is whether a person can plan a walk in such a way that he will cross each of these bridges once but not more than once.

Euler recognized the combinatorial nature of the problem and his solution can be phrased in terms of the graph in Figure 1. An **Euler path** in a graph G is a path containing each edge of G exactly once. Euler showed that such a path is possible if and only if no more than 2 vertices of G have odd local degree and G is connected (an obviously necessary condition). Hence, the answer to the Königsberg bridge problem is negative, there being 4 odd vertices. If the graph has exactly two odd vertices **a** and **b**, then an Euler path must have **a** and **b** as end vertices. If no vertices are odd, the graph contains a closed Euler path.

A constructive, algorithmic demonstration of the Euler path result borrowing from [4, p. 165] and [10, p. 39] is as follows: Suppose graph G has exactly two

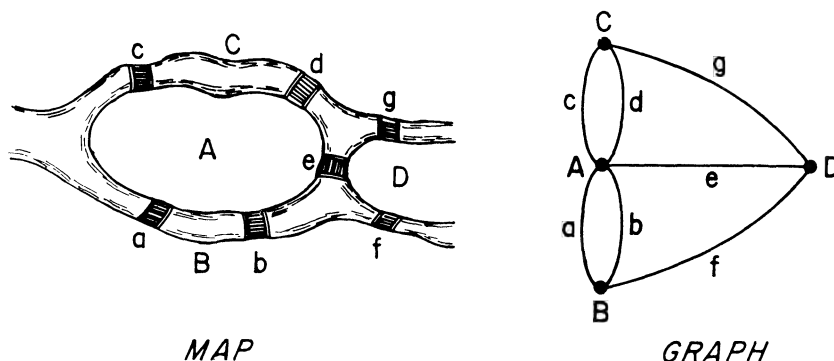


FIG. 1

odd vertices **a** and **b**. Start growing a path π_1 at vertex **a** and continue it as far as possible without repeating any edges. This path cannot get stuck at an even vertex because each time a vertex is crossed by the path two of its edges are used; furthermore, the path cannot stop at **a** because the first edge of the path is adjacent to vertex **a** leaving an even number of unused edges for subsequent crossings through **a**. Therefore, path π_1 stops at the only other odd vertex, **b**. If π_1 includes all edges of G , then it is the required Euler path. If not, we delete the edges of π_1 from G and obtain a new graph G' whose local degrees are all even since π_1 meets each vertex through an even number (possibly zero) of edges with the exception of **a** and **b**. The connectedness of G implies there is a vertex **c** on π_1 which is contained in an edge of G' . We construct a path π_2 in G' starting at **c**, which must return to **c** because all local degrees of G' are even. Enlarge π_1 to the "spliced" path $\pi_1(\mathbf{a}, \mathbf{c})/\pi_2/\pi_1(\mathbf{c}, \mathbf{b})$ and repeat the process until π_1 contains all edges of G . It can be shown by a similar argument [10, p. 40] that a graph with $2N > 0$ odd vertices can be covered by exactly N paths. More detailed discussions on Euler paths can be found in Ore [10] and Berge [4].

4. Edmonds' lemma. Edmonds [6, p. 453] gave a one-sentence proof of the theorem of Berge [1] based on the following lemma:

LEMMA E. *Let M_1 and M_2 be matchings in graph G and let $M_1 + M_2$ denote the set of edges in M_1 or M_2 but not both. Then the subgraph G_{12} formed by $M_1 + M_2$ has connected components which are paths and circuits, each of which is M_1 -alternating as well as M_2 -alternating. Each end vertex of these paths is either M_1 -exposed or M_2 -exposed.*

Proof. No vertex of G_{12} has local degree greater than two since a vertex can meet at most one M_1 edge and one M_2 edge. Hence, the graph G_{12} consists entirely of paths and circuits. Let **a** be an end vertex of one of the paths in G_{12} and suppose the adjacent end edge belongs to $M_1 \cap \bar{M}_2$. Since M_1 is a matching, **a** is not adjacent

to any other M_1 edge. Any M_2 edge adjacent to a would therefore be in $M_1 + M_2$ contradicting the fact that a is an end vertex of G_{12} . We conclude that a is M_2 -exposed.

To prove a non-maximum matching M contains an M -augmenting path is now a matter of simple arithmetic! If M' is a matching larger than M , then some component of the subgraph $M + M'$ must contain more M' edges than M edges implying the end edges are in \bar{M} and M -exposed.

The natural decomposition of $M_1 + M_2$ into alternating paths depends heavily on the special "oneness" of matchings (1-packings). An analogous result for general δ -packings requires some extra device for generating the alternating paths. The Euler paths of Section 3 supply an adequate mechanism for this purpose.

5. Alternating Euler paths. We begin this section by applying the Euler path idea to prove a generalization of Edmonds' Lemma E. We have used the notation $d(v)$ to represent the local degree of vertex v in graph G . In what follows, we shall be concerned with the various local degrees of a single vertex v in various subgraphs H_i of G and we use $d(v, H_i)$ to denote the number of edges of H_i which contain v .

LEMMA EEZ. *Let H_1 and H_2 be subgraphs of G which have no common edges. Then the subgraph G_{12} generated by $H_1 \cup H_2$ can be decomposed into an edgewise disjoint family of paths whose edges alternate between H_1 and H_2 , such that each path is one of the following types:*

- (1) *Closed paths of even length.*
- (2) *Closed paths of odd length such that the unique vertex v incident to two adjacent edges of the same H_i satisfies $d(v, H_i) - d(v, H_j) \geq 2$ where $j \neq i$.*
- (3) *Non-closed paths such that if e is an end edge in H_i containing end vertex v , then $d(v, H_i) - d(v, H_j) \geq 1$, where $j \neq i$.*

Proof. Let $\Delta_{12}(v) = d(v, H_1) - d(v, H_2)$ for all vertices in G_{12} and call vertex v **balanced**, **positive** or **negative** according as $\Delta_{12}(v)$ is zero, positive or negative. If all vertices of G_{12} are balanced, then each connected component of G_{12} enjoys the same property and hence contains an **Euler circuit** (a closed Euler path). Furthermore, since $d(v, H_1) = d(v, H_2)$ for each vertex, the Euler path can be chosen to alternate between edges of H_1 and H_2 , and so becomes a path of type 1.

If G_{12} contains unbalanced vertices, let v_1 be one such and, for convenience, suppose it to be positive. The argument is similar for negative vertices.

Since v_1 is positive $d(v_1, H_1) - d(v_1, H_2) \geq 1$ and, therefore, $d(v_1, H_1) \geq 1$. Let e_1 be one of the edges of H_1 adjacent to v_1 and let v_2 be the other vertex of e_1 . We select an edge e_2 among the H_2 edges at v_2 and add it to the path (if there exists such an edge). This process is continued as long as the path alternates between H_1 and H_2 and does not use the same edge twice. The finiteness of the graph ensures termination; this can happen in several ways.

If the path terminates at v_i via edge e_{2p} in H_2 , then $v_i \neq v_1$ since the positiveness

of v_1 ensures that every time we enter that vertex via an H_2 edge, there will be an unused H_1 edge available for exit. Since the path uses up equal numbers of H_1 and H_2 edges at each vertex it crosses (except v_1 and v_i), we can be forced to stop at v_i after edge e_{2p} in H_2 only if v_i is negative. In this case, the path $\{e_1, e_2, \dots, e_{2p}\}$ is of type 3. Deleting this path from G_{12} produces a new graph G'_{12} in which the path tracing can be resumed. Had v_1 been a negative vertex, the even length path would have ended at a positive vertex v_i . In either case, the deletion of such a path may create some balanced vertices but never any positive or negative ones.

If the path from positive v_1 ends at v_i via edge e_{2p+1} in H_1 , then either $v_i = v_1$ and $\Delta_{12}(v_1) \geq 2$ or else $v_i \neq v_1$ and v_i is positive. The first case gives a path of type 2 and the second case one of type 3. Similar results hold if v_1 is negative and once again the paths can be deleted and the path tracing resumed in the reduced graph. When we arrive at a reduced graph with no unbalanced vertices, we decompose each connected component into a path of type 2, as indicated earlier in the proof. The path tracing must terminate for lack of edges or unbalanced vertices so the lemma is proved.

We call the paths of Lemma EEZ **alternating Euler paths**.

COROLLARY 1. *If α_1 is the number of odd-length paths (possibly closed) with more H_1 edges than H_2 , and similarly for α_2 , then*

$$\alpha_1 - \alpha_2 = |H_1| - |H_2| = \frac{1}{2} \sum_{v \in G_{12}} \Delta_{12}(v).$$

Proof. Even-length alternating paths have equal numbers of edges from H_1 and H_2 and so contribute nothing to the expression $|H_1| - |H_2|$. Each odd-length alternating path has exactly one more H_1 edge than it has H_2 edges or vice versa thereby contributing ± 1 to $|H_1| - |H_2|$. Because each edge is counted twice when local vertex degrees are summed over all vertices

$$\sum_{v \in G_{12}} \Delta_{12}(v) = \sum_{v \in G_{12}} d(v, H_1) - \sum_{v \in G_{12}} d(v, H_2) = 2|H_1| - 2|H_2|.$$

COROLLARY 2. *Let H_1 and H_2 be as in Lemma EEZ and let $|H_1| > |H_2|$. Then G_{12} contains a path π which alternates between H_1 and H_2 has end edges in H_1 and end vertices v_1 and v_2 satisfying one of the following conditions:*

- (1) $\Delta_{12}(v_i) \geq 1$ for $i = 1, 2$ if $v_1 \neq v_2$.
- (2) $\Delta_{12}(v_i) \geq 2$ if $v_1 = v_2$.

Proof. By Corollary 1 we get $\alpha_1 - \alpha_2 = |H_1| - |H_2| \geq 1$ and hence $\alpha_1 \geq 1$. This assures the existence of an odd length path of type 2 or 3 with end edges in H_1 .

6. The theorem of Berge-Norman-Rabin. We can now give a simple proof of the Berge-Norman-Rabin theorem [4, p.175] using Lemma EEZ and Corollary 2.

THEOREM BNR. *If P is a non-maximum δ -packing in graph G , then G contains a (P, δ) -augmenting path.*

Proof. Let P^* be a larger δ -packing and put $H_1 = P^* - P$ and $H_2 = P - P^*$. Applying Lemma EEZ and Corollary 2 (since $|H_1| > |H_2|$), we get a path which alternates edges of P and \bar{P} (i.e., P -alternating), has end edges in \bar{P} and end vertices v_1 and v_2 each satisfying condition (1) or (2) of Corollary 2. Since the edges in $P \cap P^*$ contribute to both terms of the expression $d(v_i, P^*) - d(v_i, P)$, we see easily that

$$d(v_i, P^*) - d(v_i, P) = d(v_i, H_1) - d(v_i, H_2).$$

Combining this with conditions (1) and (2) of Corollary 2 and the inequality $d(v_i, P^*) \leq \delta(v_i)$, we find that

$$d(v_i, P) \leq \delta(v_i) - 1 \quad \text{for } i = 1, 2 \quad \text{if } v_1 \neq v_2$$

$$d(v_1, P) \leq \delta(v_1) - 2 \quad \text{if } v_1 = v_2.$$

Hence, the path is (P, δ) -augmenting.

7. Graphs with edge dichotomies. Any graph G whose edge set E has been dichotomized (i.e., partitioned into two subsets) can be decomposed by Lemma EEZ into a family of edge-disjoint alternating Euler paths with fairly natural conditions on the end vertices. If $E = E_1 \cup E_2$ is the dichotomy, let H_i for $i = 1, 2$ be the subgraph of G generated by the edge set E_i , and apply Lemma EEZ (in this case $G_{12} = G$). We separated Lemma EEZ from the proof of Theorem BNR because the alternating path decomposition is a general phenomenon not dependent on packings or covers or local degree constraints. The following corollaries strengthen Lemma EEZ somewhat:

COROLLARY 3. *If graph G_{12} in Lemma EEZ is connected, then it can be decomposed so that there is, at most, one path of type 1, and that only if it is the sole path covering all of G_{12} .*

Proof. Because G_{12} is connected, the even-length closed (type 1) paths can be “spliced” together or into other paths of types 2 or 3. If at least one path of type 2 or 3 exists, then all the paths of type 1 can be made to disappear into one or more of the paths of types 2 or 3. The splicing is similar to that used in section 2 for Euler paths. Clearly, at least one path is required, so a single type 1 path is possible.

To characterize further the alternating path decompositions, we need some additional terminology. Let α_0 be the number of even-length paths of type 3 and let α_i for $i = 1$ and 2 be, as before, the number of odd-length paths of types 2 and 3 with more H_i edges. It is then obvious that the number of paths of type 2 or 3 is exactly $(\alpha_0 + \alpha_1 + \alpha_2)$. We call $\Delta^T = \sum |\Delta_{12}(v)|$ the **total vertex imbalance** for dichotomy (H_1, H_2) .

COROLLARY 4. *The path decomposition of a connected G_{12} as presented in Corollary 3, is minimum in the sense that no other representation of G_{12} as a family of alternating paths has fewer paths. Furthermore, the number of paths of type 2 or*

3 is related directly to the vertex differentials $\Delta_{12}(v)$ by

$$\alpha_0 + \alpha_1 + \alpha_2 = \frac{1}{2} \sum_{v \in G_{12}} |\Delta_{12}(v)| = \Delta^T/2.$$

Proof. First we show that $\Delta^T/2$ is a lower bound for the number of paths in an alternating family. Let \mathcal{F} be a family of alternating paths for G_{12} and consider a single vertex v with differential $\Delta_{12}(v) > 0$. The edges of H_1 and H_2 incident to v can be paired off except for exactly $|\Delta_{12}(v)|$ extra H_1 edges. Each pair of edges corresponds to the occurrence of v as an internal vertex of an alternating path of \mathcal{F} . Each extra edge must represent the occurrence of v as an end vertex. An identical argument holds for $\Delta_{12}(v) < 0$. The paths of any alternating path decomposition must hence account for at least Δ^T end vertex occurrences, but each path can handle, at most 2 so $\Delta^T/2$ is indeed a lower bound. In the proof of Lemma EEZ, paths of type 2 or 3 are constructed only between end vertices which are currently unbalanced and each path deletion decreases the total vertex imbalance by 2 units. The construction of type 2 and type 3 paths terminates when the total vertex imbalance is reduced to zero so the total number of such paths is precisely $\Delta^T/2$. This establishes the formula and the minimality follows because our particular decomposition achieves the lower bound.

Let us call $\Delta^\Sigma = \sum \Delta_{12}(v)$ the **net vertex imbalance**. It is then tempting to ask if a decomposition of G_{12} into alternating paths can be accomplished with $\alpha_1 + \alpha_2 = |\Delta^\Sigma|/2$, it being clear that $\alpha_1 + \alpha_2 \geq |\Delta^\Sigma|/2$. If equality does hold, then either $\alpha_1 = \Delta^\Sigma/2$ while $\alpha_2 = 0$, or else $\alpha_2 = -\Delta^\Sigma/2$ while $\alpha_1 = 0$. Figure 2 depicts a simple dichotomized graph with $\Delta^\Sigma = 0$ which requires $\alpha_1 + \alpha_2 \geq 2$.

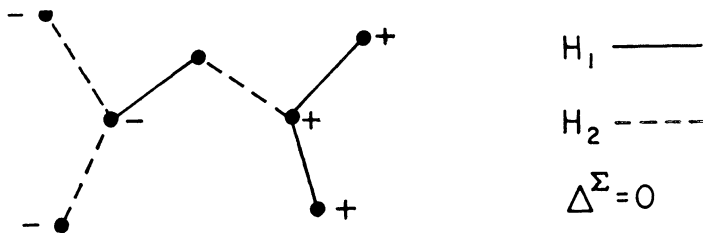


FIG. 2

On the other hand, this seems to result from the lack of connectedness between positive and negative vertices so equality may be achievable under some sort of multiple connectedness assumption. In any case, it would be interesting to find a decomposition with minimum value of $\alpha_1 + \alpha_2$ and know how the minimum relates to the structure of graph G_{12} .

8. Acknowledgements. I am deeply indebted to former colleagues at the National Bureau of Standards — J. Edmonds, A. J. Goldman, and C. Witzgall. In particular, the ideas and style of J. Edmonds in the area of combinatorial graph theory have been a great influence on my approach to

such problems. The work reported here was begun while the author was at the National Bureau of Standards and was partially supported by the Army Research Office (Durham).

More recent support has come from the National Science Foundation and the Stanford Linear Accelerator Center (funded by the Atomic Energy Commission).

References

1. C. Berge, Two theorems in graph theory, Proc. Nat. Acad. Sci. U. S. A., 43 (1957) 842–844.
2. R. Z. Norman and M. O. Rabin, An algorithm for a minimum cover of a graph, Proc. Amer. Math. Soc., 10(1959) 315–319.
3. J. Edmonds, Covers and packings in a family of sets, Bull. Amer. Math. Soc., 68 (1962) 494–499.
4. C. Berge, The Theory of Graphs and its Applications, Methuen, London, 1962.
5. A. J. Goldman, Optimal matchings and degree-constrained subgraphs, J. Res. Nat. Bur. Stds., 68B (1964) 27–29.
6. J. Edmonds, Paths, trees, and flowers, Canad. Math. J., 17 (1965) 449–467.
7. L. Euler, The seven bridges of Königsberg, The World of Mathematics, J. R. Newman ed., Simon and Schuster, New York, 1956, pp. 573–580.
8. C. Witzgall and C. T. Zahn Jr., Modification of Edmonds' maximum matching algorithm, J. Res. Nat. Bur. Stds., 69B (1965) 91–98.
9. J. Edmonds, Maximum matching and a polyhedron with 0, 1-vertices, J. Res. Nat. Bur. Stds., 69B (1965) 125–130.
10. O. Ore, Theory of Graphs, Amer. Math. Soc., Providence, Rhode Island, 1962.
11. J. Petersen, Die Theorie der regulären Graphen, Acta Math., 15 (1891) 193–220.
12. W. T. Tutte, The factors of graphs, Canad. Math. J., 4 (1952) 314–328.

MATHEMATICAL NOTES

EDITED BY ROBERT GILMER

The present backlog for this Department is substantial. Until further notice, new manuscripts cannot be accepted. This moratorium will probably continue until June 1, 1973; authors are requested to hold their manuscripts pending a further announcement.

STABLE LAWS AND THE IMBEDDING OF L^p SPACES

MAREK KANTER, Tulane University, New Orleans

1. Introduction. It was conjectured by Banach [1] that if $1 \leq p \leq q \leq 2$ then there exists a linear isometry from $L^q[0, 1]$ into $L^p[0, 1]$. The truth of this conjecture is now known (see [6]).

In fact, more is known. In [3], Lemma 1 on p. 238, there is given a proof which essentially demonstrates that $L^q[0, 1]$ can be linearly and isometrically imbedded into $L^p[0, 1]$ if $0 < p \leq q \leq 2$. However, in Theorem 2 of [3] on p. 238, when this Lemma is applied, p and q are restricted to be equal to or greater than 1. Also in an article as recent as [7], the solution of Banach's conjecture when $0 < p \leq q \leq 2$ is left as an open problem.

We have decided that it would be useful to write an expository and reasonably self-contained paper that proves the truth of Banach's conjecture for $0 < p \leq q \leq 2$. Our proof is different, and more direct, than the proof in [3]. It rests completely upon our being able to define a stochastic integral with respect to a certain stochastic process. The stochastic integral that we define is an interesting object in itself, and one of the purposes of this paper is to introduce this stochastic integral to probabilists and analysts.

2. Notations and definitions. A measure space is a triple $(\Omega, \mathcal{B}, \mu)$, where Ω is any set, \mathcal{B} is a sigma-field of subsets of Ω , and μ is a nonnegative measure on \mathcal{B} . We assume from now on that all our measure spaces are complete, i.e., if $C \subset D$ with $D \in \mathcal{B}$ and $\mu(D) = 0$ then $C \in \mathcal{B}$.

For $p > 0$ we shall write $L^p[\Omega]$ to denote the set of all real valued \mathcal{B} measurable functions with

$$\|f\|_p = \left[\int_{\Omega} |f|^p d\mu \right]^{1/p} < \infty.$$

If $(\Omega', \mathcal{B}', \mu')$ is another measure space, then a linear mapping T from $L^{p'}[\Omega']$ to $L^p[\Omega]$ is said to be an isometry if $\|T(f)\|_p = \|f\|_{p'}$ for $f \in L^{p'}[\Omega']$.

3. Some concepts from probability. If μ is a probability measure, i.e., if $\mu(\Omega) = 1$, then we use the symbol P instead of μ , and we call (Ω, \mathcal{B}, P) a probability triple. By definition a random variable on Ω is a real valued \mathcal{B} measurable function on Ω , and a stochastic process on Ω is a collection of random variables on Ω . If X and Y are two random variables on Ω then we say that $X = Y$ a.s. if $P\{w \mid w \in \Omega, X(w) \neq Y(w)\} = 0$.

If X is a random variable on Ω then the characteristic function $\phi_X(v)$ is defined for all real v by

$$\phi_X(v) = E(e^{ivX}),$$

where E stands for integration with respect to the measure P . (In the following we shall sometimes write $\exp(z)$ instead of e^z to denote exponentiation.)

If X is a random variable on Ω we say that X is *symmetric stable of index q* ($q > 0$) if for all real v ,

$$\phi_X(v) = \exp(-k|v|^q)$$

for some $k \geq 0$. (We also say that X is symmetric q -stable.) We notice as in [4, p. 486] that if $q > 2$ then there is no random variable X such that $\phi_X(v) = \exp(-k|v|^q)$, because

$$E(X^2) = -\frac{d^2}{dv^2} \phi_X(0) = 0$$

for $q > 2$.

We say that a collection (X_1, \dots, X_n) of random variables on Ω , is independent if

for all v_1, \dots, v_n, v real we have $\phi_Z(v) = \prod_{i=1}^n \phi_{X_i}(v_i v)$, where $Z = \sum_{i=1}^n v_i X_i$. This is not the usual definition of independent random variables, but it is certainly equivalent to it, (see [4], p. 495). From this definition we conclude that if each of the above X_i is symmetric q -stable with $\phi_{X_i}(v) = \exp(-k_i |v|^q)$ for $i \in [1, \dots, n]$, then $Z = \sum_{i=1}^n v_i X_i$ is also symmetric stable of index q and in fact

$$(*) \quad \phi_Z(v) = \exp \left(- \left(\sum_{i=1}^n |v_i|^q k_i \right) |v|^q \right).$$

We end this introduction with a few brief remarks about convergence in measure. Namely, we say that a sequence X_n of random variables converges in measure to the random variable X if for all $\varepsilon > 0$ we have that $\lim_{n \rightarrow \infty} P[|X_n - X| \geq \varepsilon] = 0$. It is not hard to see that this is equivalent to having

$$\lim_{n \rightarrow \infty} E(|X_n - X|/1 + |X_n - X|) = 0.$$

Hence we conclude that if we define $\rho(X, Y)$ to be $E(|X - Y|/1 + |X - Y|)$ and if we identify random variables that are a.s. equal, then ρ makes the set of random variables on a probability space (Ω, \mathcal{B}, P) into a metric space M , and in fact ρ metrizes the notion of convergence in measure.

4. Existence of stable processes. We fix q in the interval $(0, 2]$. Paul Levy not only discovered q -stable random variables, but he also proved the existence of a stochastic process $(X(t) | t \in [0, 1])$ on a probability space (Ω, \mathcal{B}, P) such that

1. $X(0) = 0$ a.s.
2. For all $0 \leq t_1 < \dots < t_n \leq 1$
the set $(X(t_{i+1}) - X(t_i) | i = 0, \dots, n-1)$ is a collection of n independent random variables such that

$$(**) \quad \phi_Z(v) = \exp(-|t-s| |v|^q)$$

for all real v and $t, s \in [0, 1]$ with $Z = X(t) - X(s)$.

In Levy's work Ω is a specific function space, but what concerns us here is only the fact that (Ω, \mathcal{B}, P) is a *separable* measure space in the measure theoretic sense. (See [5], p. 168.)

We refer the interested reader to Breiman's book [2, Chapter 14] for an account of the existence of such processes. Breiman (and Levy as well) prove the existence of more general processes. (Namely $(**)$ is replaced by the condition that $\phi_Z(v) = \phi_{|t-s|}(v)$, i.e., the functions ϕ_Z depend only on $|t-s|$.)

5. The main theorem. Suppose we are given $(X(t) | t \in [0, 1])$ a stochastic process satisfying conditions 1. and 2. above. If f is a step function, i.e.,

$$f = \sum_{i=1}^n c_i I_{[t_i, t_{i+1})}$$

where c_1, \dots, c_n are real numbers, $0 \leq t_1 < \dots < t_{n+1} \leq 1$, and $I_{[t_i, t_{i+1})}$ is the indicator function of the half open interval $[t_i, t_{i+1})$ then we set

$$\int f dX = \sum_1^n c_i (X(t_{i+1}) - X(t_i)).$$

We call $\int f dX$ the *stochastic integral* of f with respect to X .

THEOREM. *There exists a linear map $f \rightarrow \int f dX$ from $L^q[0,1]$ into the set of symmetric q -stable random variables on (Ω, \mathcal{B}, P) which agrees with the definition of the stochastic integral on step functions and which satisfies*

(***) $\phi_Z(v) = \exp(-(\|f\|_q^q |v|^q))$ for $f \in L^q[0,1]$, v real, and $Z = \int f dX$.

Proof. (***) holds for step functions because of (*) and (**). Also it is trivial to see that the map $f \rightarrow \int f dX$ is linear for step functions. Now for any $f \in L^q[0,1]$ there exists a sequence of step functions f_n such that $\|f - f_n\|_q \rightarrow 0$. In particular

$$\lim_{n,m \rightarrow \infty} \|f_n - f_m\|_q \rightarrow 0.$$

Now we define $Z_{n,m} = \int f_n dX - \int f_m dX$. Also we define ϕ_n to be the characteristic function of $\int f_n dX$ while $\phi_{n,m}$ is to be the characteristic function of $Z_{n,m}$.

By Breiman [2, p. 171] we have that for any $\varepsilon > 0$

$$P[|Z_{n,m}| \geq \varepsilon] \leq k_0 \varepsilon \int_0^{1/\varepsilon} (1 - \phi_{n,m}(v)) dv,$$

where k_0 is a positive constant. If we use (***) to substitute for $\phi_{n,m}$ in this inequality we conclude that $\lim_{n,m \rightarrow \infty} P[|Z_{n,m}| \geq \varepsilon] = 0$.

This implies that $\int f_n dX$ is a Cauchy sequence in the metric of convergence in measure. Now it follows from Halmos [5, Theorem E, p. 93] that there exists a random variable which is unique up to sets of measure zero (i.e., it defines a unique element of M) and which is the limit of $\int f_n dX$ in this metric. We call this random variable $\int f dX$. Furthermore by Lebesgue's bounded convergence theorem (see Halmos [5], Theorem D, p. 110 and note that it works equally well for complex valued functions) we conclude that if ϕ_Z is the characteristic function of $\int f dX$ then $\phi_Z(v) = \lim_{n \rightarrow \infty} \phi_n(v)$, and from this we conclude that (***) is valid for $f \in L^q[0,1]$.

To finish we prove the linearity of the mapping $f \rightarrow \int f dX$. This mapping is clearly linear on step functions. Now let a, b be two real numbers. Then

$$\int (af + bg) dX = a \int f dX + b \int g dX \text{ a.s.}$$

upon taking limits along appropriate sequences of step functions f_n and g_n . This completes the proof of the main theorem.

COROLLARY. *If $0 < p < q \leq 2$ then there exists a linear isometry from $L^q[0,1]$ to $L^p[0,1]$.*

Proof. We first show that if X is symmetric stable of index q then $E(|X|^p) < \infty$ if $0 < p < q$. This is a problem in Feller [4, p. 215]. We notice as he does that from the symmetrization inequalities in [4, p. 147, 148] we can get

$$\frac{1}{2}(1 - \exp\{-nP[|X| \geq n^{1/q}t]\}) \leq P[|X| \geq t]$$

for any natural number n and any positive number t . If we set $t = n^\varepsilon$ with $\varepsilon > 0$, we conclude $nP[|X| \geq n^{1/q+\varepsilon}]$ is bounded. This easily implies that $n^{1+\varepsilon}P[|X| \geq n^{(1+\varepsilon)/q+\varepsilon}]$ is also bounded. If we choose ε so small that $p < \{(1+\varepsilon)/q + \varepsilon\}^{-1}$ then we conclude that $\sum_1^\infty P[|X|^p \geq n] < \infty$, and hence that $E(|X|^p) < \infty$.

Let us write $C_{q,p} = E(|X_0|^p)^{1/p}$, where $\phi_{X_0}(v) = e^{-|v|^q}$. Then if $\phi_Y(v) = e^{-k|v|^q}$ it follows that $k^{1/q}C_{q,p} = (E(|Y|^p))^{1/p}$ since Y is distributed like $(k)^{1/q}X$.

Now let $(X(t))_{t \in [0,1]}$ be a stochastic process on (Ω, \mathcal{B}, P) which satisfies Conditions 1 and 2 above. By the theorem there exists a linear mapping $f \rightarrow \int f dX$ from $L^q[0,1]$ into the set of symmetric q -stable random variables on (Ω, \mathcal{B}, P) . Furthermore we have just verified that the mapping $f \rightarrow C_{q,p}^{(-1)} \int f dX$ is a linear isometry from $L^q[0,1]$ into $L^p[\Omega]$.

To finish the proof of the corollary we must map $L^p[\Omega]$ linearly and isometrically into $L^p[0,1]$. For those who are willing to believe, I can simply say that we could have taken (Ω, \mathcal{B}, P) to be the unit interval with Lebesgue measure. However, those readers who have referred to Breiman for the existence of the process $X(t)$ have in mind the specific separable measure space (Ω, \mathcal{B}, P) constructed there (see [2, p. 306]). These readers must now check that (Ω, \mathcal{B}, P) is nonatomic as well as separable, and then apply Theorem C of [4, p. 173] to conclude that the measure algebra of (Ω, \mathcal{B}, P) is isomorphic to the measure algebra of the unit interval, which of course implies that $L^p[\Omega]$ is linearly isometric with $L^p[0,1]$. (Another route would be, instead of checking that (Ω, \mathcal{B}, P) is nonatomic, to generalize Theorem C of [4, p. 173] so that the hypothesis of nonatomicity is dropped. In this case one no longer gets a measure algebra isomorphism into the measure algebra of the unit interval, but rather just a homomorphism. This implies that $L^p[\Omega]$ can be mapped linearly and isometrically into $L^p[0,1]$.)

References

1. S. Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932.
2. L. Breiman, *Probability*, Addison-Wesley, Reading, Mass. 1968.
3. J. Bretagnolle, D. Dacunha-Castelle, and J. L. Krivine, Lois stables et espaces L^p , *Ann. Inst. H. Poincaré Sect. B*, 2 (1966) 231–259.
4. W. Feller, *An Introduction to Probability Theory and its Applications*, vol. 2, Wiley, New York, 1966.
5. P. Halmos, *Measure Theory*, Van Nostrand, New York, 1950.
6. J. Lindenstrauss and A. Pełczyński, Absolutely summing operations in L_p spaces and their applications, *Studia Math.*, 29 (1968) 275–326.
7. W. J. Stiles, On properties of subspaces of l_p , $0 < p < 1$, *Trans. Amer. Math. Soc.*, 149 (1970) 405–416.

A CONVEX MATRIX FUNCTION

M. H. MOORE, University of Florida

Let P_n be the class of all (strictly) positive definite symmetric $n \times n$ matrices, and let N_n be the class of all non-negative definite symmetric $n \times n$ matrices. Thus $P_n \subset N_n$.

We use the standard notation " $A \geq B$ " to indicate that $A - B \in N_n$, and similarly " $A > B$ " means that $A - B \in P_n$.

Our objective is to state and prove the following theorem:

THEOREM. *Let A and B belong to P_n , and let $0 \leq \lambda \leq 1$. Then*

$$[\lambda A + (1 - \lambda)B]^{-1} \leq \lambda A^{-1} + (1 - \lambda)B^{-1}.$$

With the obvious definitions, the theorem may be interpreted as saying that the matrix inverse function is convex on P_n .

The theorem complements a recent result of Ky Fan [4] who has obtained the convexity property of the theorem for the so-called " M matrices" of Ostrowski.

Proof. Since A and B are symmetric and positive definite, the matrices are simultaneously diagonalizable. More precisely, the following well-known (see, for example, [5], pp. 310) result holds:

LEMMA. *Let A and B be symmetric $n \times n$ matrices with $A \in P_n$. Then there exists a (real) non-singular matrix Q such that*

$$A = QQ^T, \quad B = QDQ^T,$$

where D is a diagonal matrix whose diagonal elements λ_i are real and are the solutions to $\det(B - \lambda A) = 0$. If, moreover, B is non-negative (positive) definite, then the λ_i are non-negative (positive).

Applying this result in our case, we have $A = QQ^T$, $B = QDQ^T$, and

$$\lambda A + (1 - \lambda)B = Q[\lambda I + (1 - \lambda)D]Q^T = QE_\lambda Q^T,$$

where Q is a non-singular matrix and $D = \text{diag}(\lambda_1, \dots, \lambda_n)$, ($\lambda_i > 0$ for $1 \leq i \leq n$), and

$$E_\lambda = \lambda I + (1 - \lambda)D.$$

Notice that D and E_λ are non-singular, all diagonal elements of each matrix being positive. Put $P = Q^{-1}$, then

$$A^{-1} = (QQ^T)^{-1} = (Q^T)^{-1}Q^{-1} = (Q^{-1})^T Q^{-1} = P^T P,$$

$$B^{-1} = P^T D^{-1} P, \text{ and } [\lambda A + (1 - \lambda)B]^{-1} = P^T E_\lambda^{-1} P.$$

We now find that

$$\begin{aligned}
\lambda A^{-1} + (1 - \lambda)B^{-1} - [\lambda A + (1 - \lambda)B]^{-1} &= \lambda P^T P + (1 - \lambda)P^T D^{-1} P - P^T E_\lambda^{-1} P \\
&= P^T [\lambda I + (1 - \lambda)D^{-1} - E_\lambda^{-1}] P \\
&= P^T R_\lambda P,
\end{aligned}$$

where $R_\lambda = \lambda I + (1 - \lambda)D^{-1} - E_\lambda^{-1}$ is a diagonal matrix. But the ii th element of R_λ is

$$(R_\lambda)_{ii} = \lambda + (1 - \lambda) \frac{1}{\lambda_i} - \frac{1}{\lambda + (1 - \lambda)\lambda_i}$$

and, since the λ_i are all positive, the convexity of the function $\phi(x) = 1/x$ for $x > 0$ shows that each $(R_\lambda)_{ii}$ is non-negative (in fact, positive unless $\lambda_i = 1$ or, trivially, $\lambda = 0$ or 1). Hence $R_\lambda \in N_n$. From this, it is clear that

$$0 \leq P^T R_\lambda P = \lambda A^{-1} + (1 - \lambda)B^{-1} - [\lambda A + (1 - \lambda)B]^{-1}$$

which is the desired result.

The matter of convexity for functions of matrices has been introduced and discussed by Krauss [6] in 1936. More recent results have been obtained by Bendat and Sherman [2] and Chandler Davis [3]. The related idea of monotone functions of matrices, which has found wide application in recent years, was brought forward by Löwner [7] in 1934. Bellman [1] page 111, gives an account of other literature in this intriguing area.

References

1. R. Bellman, *Introduction to Matrix Analysis* (2nd ed.), McGraw-Hill, New York, 1970.
2. J. Bendat and S. Sherman, Monotone and convex operator functions, *Trans. Amer. Math. Soc.*, 79 (1955) 58-71.
3. C. Davis, Notions generalizing convexity for functions defined on spaces of matrices, *Proc. Symposia Pure Math.*, vol. 7: Convexity, American Mathematical Society, (1963) 187-201.
4. Ky Fan, Inequalities for the sum of two M -matrices, *Inequalities*, Oved Shisha ed., Academic Press, New York, 1967, pp. 105-117.
5. F. R. Gantmacher, *Matrix Theory*, vol. 1, Chelsea, New York, 1959.
6. F. Krauss, Über konvexe Matrixfunktionen, *Math. Z.*, 41 (1936) 18-42.
7. K. Löwner, Über monotone Matrixfunktionen, *Math. Z.*, 38 (1934) 177-216.

SOLUTION OF FEJES TÓTH'S ILLUMINATION PROBLEM

B. R. HENRY, Syracuse University

In this note we complete the disproof of Fejes Tóth's illumination conjecture ([1] 1970) using the illumination function of A. Heppes (Guy and Klee 1971, [2] p. 1118): $f(x)$ equals 1, $\frac{1}{2}(m+1) - x$, or 0, according as $0 \leq x < \frac{1}{2}(m-1)$, $\frac{1}{2}(m-1) \leq x < \frac{1}{2}(m+1)$, or $\frac{1}{2}(m+1) \leq x$.

Let m be transcendental and suppose a lamp L_0 stands at the origin. If illumi-

nation is to be uniform a lamp L_1 must stand at m or at -1 in order to cancel the kink in f at $\frac{1}{2}(m-1)$. Let the position of L_1 be ξ_1 . To cancel the corresponding kink in L_1 a lamp L_2 must be placed at $\xi_1 + m$ or at $\xi_1 - 1$; call its position ξ_2 . Continuing, an infinite sequence $0, \xi_1, \xi_2, \xi_3, \dots$ of lamp coordinates is generated (by irrationality of m) that are required if illumination is to be constant.

No two of the ξ 's differ by a multiple of $m/(m+1)$. For if k is an integer and

$$\xi_i - \xi_j = \frac{km}{m+1},$$

then there are integers a, b such that

$$am + b = \frac{km}{m+1},$$

so $am(m+1) + b(m+1) - km = 0$ or $am^2 + (a+b-k)m + b = 0$. By transcendentality of m , $a = b = (a+b-k) = 0 \Rightarrow k = 0$.

If lamps are erected in a finite union of congruent point lattices at average density $(m+1)/m$ then there are for some k , k lattices of span $km/(m+1)$. In any subset of more than k coordinates, some two coordinates must differ by a multiple of $km/(m+1)$. By the remarks above no such arrangement gives uniform illumination.

Research supported by NSF grant GP-31379.

References

1. L. Fejes Tóth, A problem of illumination, this MONTHLY, 77 (1970) 869-870.
2. R. Guy and V. Klee, Monthly research problems, 1969-71, this MONTHLY, 78 (1971) 1113-1122.

A COVERING THEOREM

J. C. KIEFFER, University of Missouri-Rolla

Let G be a nonempty bounded open set in E^m , Euclidean m -space. It is not possible to find countably many open balls B_i , each contained in G , such that $G = \limsup B_i$, and $\sum_i m(B_i) < \infty$. (Here, m denotes Lebesgue measure.) This follows from the well-known result from measure theory that if $\sum_i m(B_i) < \infty$, then $m(\limsup B_i) = 0$. It is possible, however, to find balls B_i in G such that $G = \limsup B_i$, and $\sum_i m(B_i)^p < \infty$, for every $p > 1$.

Proof. $B(x, r)$ denotes the open ball with center x and radius r . For each positive integer n , select all balls of form $B(x/n, r_n)$ such that:

$$(1) \quad B\left(\frac{x}{n}, r_n\right) \subset G;$$

$$(2) \quad x \text{ is a lattice point}$$

(that is, its coordinates are all integers); and

$$(3) \quad r_n = \frac{\sqrt{m}}{n^{1+(1/m)}}.$$

For each n , the number of balls selected is $O(n^m)$, since G is bounded. The total collection of balls selected, as n ranges over all positive integers, gives us a countable collection of balls B_i . For an appropriate constant C , we have

$$\sum_i m(B_i)^p \leq C \sum_n n^m r_n^{mp}, \quad p > 1.$$

Substituting in the value for r_n , we see that $\sum_i m(B_i)^p < \infty$.

To conclude the proof, we show that $G = \limsup B_i$. Suppose then that $y \in G$. For sufficiently large n , we have $B(y, 2r_n) \subset G$. We use now a theorem which allows us to approximate points in E^m by points with rational coordinates in a certain way (see [1], Theorem 4.6). This theorem states that for infinitely many n , it is possible to find a lattice point x_n such that $y \in B(x_n/n, r_n)$. For infinitely many n , then, $B(y, 2r_n) \subset G$ and $y \in B(x_n/n, r_n)$ hold simultaneously. For each such n , a simple application of the triangle inequality shows that $B(x_n/n, r_n) \subset G$. From the way in which the B_i were defined we conclude that $y \in B_i$ for infinitely many i .

Reference

1. Ivan Niven, *Irrational Numbers*, MAA Carus Mathematical Monograph 11, (1956) 47.

DISTRIBUTIVITY OVER THE DIRICHLET PRODUCT AND COMPLETELY MULTIPLICATIVE ARITHMETICAL FUNCTIONS

ERIC LANGFORD, University of Maine

Recently in this MONTHLY, some interest has been expressed in the relationship between an arithmetical function's being completely multiplicative and its being distributive over certain Dirichlet products. Lambek [3] proved (Theorem 1) that the arithmetical function f is completely multiplicative if and only if it distributes over *every* Dirichlet product. Problems of Carlitz [2] and Sivaramakrishnan [4] have shown that f is necessarily completely multiplicative if and only if it distributes over certain *particular* Dirichlet products. Apostol [1] also gives various conditions involving the Dirichlet product that guarantee that f is completely multiplicative.

In this paper, we shall give sufficient conditions on a particular Dirichlet product in order that every arithmetical function which distributes over that product must necessarily be completely multiplicative. These conditions will be general enough so that the results of [1], [2], [3], and [4] in this area will follow as corollaries.

We shall say that f **distributes over** the Dirichlet product $g * h = k$ if $fg * fh = fk$.

It is immediate that if f is completely multiplicative, then f distributes over every Dirichlet product and hence over any particular such product. But every function f , multiplicative or not, which satisfies $f(1) = 1$ distributes over the product $\delta * \delta = \delta$, where $\delta(n) = 1$ if $n = 1$ and 0 otherwise. The problem is to find those Dirichlet products which are sufficiently "sensitive" so that only completely multiplicative functions will distribute over them. Carlitz's problem shows that $1 * 1 = \tau$ is such a product, where $\tau(n)$ denotes the number of divisors of n , and where $1(n) = 1$ for all n . Sivaramakrishnan's problem essentially shows that $\phi * 1 = I$ is such a product, where $I(n) = n$ is the identity function and where ϕ is Euler's totient function.

If $k = g * h$ is any Dirichlet product, we notice that

$$(1) \quad k(n) = g(1)h(n) + g(n)h(1)$$

whenever n is prime. If (1) holds only when n is prime, we say that the product $k = g * h$ is **discriminative**.

THEOREM 1. *Suppose that $f(1) \neq 0$. Then f is completely multiplicative if and only if it distributes over some discriminative product $k = g * h$.*

Proof. As already remarked, if f is completely multiplicative, then it must distribute over every Dirichlet product. Assume then that $f(1) \neq 0$ and that f distributes over the discriminative product $k = g * h$. We show first that necessarily $f(1) = 1$. Since $k = g * h$ is discriminative, it follows that $k(1) \neq 0$ for otherwise (1) would hold with $n = 1$. But

$$f(1)k(1) = fk(1) = fg(1)fh(1) = f(1)^2g(1)h(1) = f(1)^2k(1),$$

and since $f(1)k(1) \neq 0$, it follows that $f(1) = 1$.

We now show by complete induction on m that if p_1, p_2, \dots, p_m is any choice of primes (distinct or not), then

$$(2) \quad f(p_1 \cdots p_m) = f(p_1) \cdots f(p_m).$$

If $m = 1$, the proposition is trivial, so assume that $m \geq 2$, and write $n = p_1 \cdots p_m$. Using the distributive property and the inductive assumption, we see that

$$[f(p_1 \cdots p_m) - f(p_1) \cdots f(p_m)] \sum' g(d)h(n/d) = 0,$$

since the terms involving $f(1)$ cancel; the prime on the sum indicates that it is to be taken over all divisors d of n other than 1 and n . But

$$\sum' g(d)h(n/d) = k(n) - g(1)h(n) - g(n)h(1),$$

and this is non-zero since $k = g * h$ is discriminative. Therefore (2) holds and the proof is complete.

A Dirichlet product $k = g * h$ will be called **partially discriminative** if for every

prime power p^i (with $i \geq 1$), the equation

$$k(p_i) = g(1)h(p_i) + g(p_i)h(1)$$

implies that $i = 1$.

THEOREM 2. *Suppose that f is multiplicative. Then f is completely multiplicative if and only if f distributes over some partially discriminative product $k = g * h$.*

Proof. If $f(1) = 0$, then it is easily shown that $f(n) = 0$ for all n . If $f(1) \neq 0$, then $f(1) = 1$ since f is multiplicative and the proof follows by showing that $f(p^m) = f(p)^m$ for every prime power p^m , using complete induction on m .

Theorem 1 still leaves the case of $f(1) = 0$ unresolved. In general, nothing can be inferred about f in this case. For example, if we let $f = 1 - |\mu|$, where μ is the Möbius function, then f distributes over the discriminative product $\mu * 1 = \delta$, even though f is not even multiplicative. Something, however, can be salvaged: it is not hard to show that if $f(1) = 0$ and if f distributes over some (not necessarily discriminative) product $k = g * h$, where $k(n)$ never vanishes, then f must be completely multiplicative; i.e., $f(n) = 0$ for all n .

We show now how the results in [1], [2], [3], and [4] are corollaries of these two theorems.

COROLLARY 1. *Sivaramakrishnan's Problem E 2196 [4] asks us to show that $fI * f^{-1} = f\phi$ if and only if f is completely multiplicative. This is equivalent to f being distributive over $\phi * 1 = I$, and Theorem 1 applies. Sivaramakrishnan's assumption that f be multiplicative is superfluous.*

COROLLARY 2. *Carlitz's Problem E 2268 [2] asks us to show that f is completely multiplicative if and only if it distributes over $1 * 1 = \tau$. Theorem 1 again applies. This also shows Lambek's Theorem 1 [3].*

COROLLARY 3. *Apostol's Theorem 2 [1] claims that if f is multiplicative (and not identically zero), then $f^{-1} = f\mu$ if and only if f is completely multiplicative. But $f^{-1} = f\mu$ if and only if f distributes over $\mu * 1 = \delta$, and Theorem 2 applies.*

COROLLARY 4. *Apostol's Theorem 8 generalizes Sivaramakrishnan's result as follows: Suppose that G is a completely multiplicative function and that $g = G * \mu$. Suppose further if p is a prime, then $G(p^i) = 1$ only if $i = 0$. If f is multiplicative, and if $fG * f^{-1} = fg$, then f is completely multiplicative. (Remark: Apostol assumes that $G(p) \neq 1$ for every prime p , but in his proof he uses the fact that $G(p^i) = G(p)^i \neq 1$ if $i \geq 1$. Since we allow complex values, these are not equivalent assumptions.)*

We need not assume even that G is multiplicative, but only that $G(p^i) = 1$ if and only if $i = 0$. Now $fG * f^{-1} = fg$ is equivalent to f being distributive over $G = (G * \mu) * 1$ and the assumption that $G(p^i) = 1$ if and only if $i = 0$ allows Theorem 2 to apply.

I would like to thank the referee for clarifying the statements of Theorems 1 and 2.

References

1. T. M. Apostol, Some properties of completely multiplicative arithmetical functions, this MONTHLY, 78 (1971) 266-271.
2. Leonard Carlitz, Problem E 2268, this MONTHLY, 78 (1971) 1140.
3. J. Lambek, Arithmetical functions and distributivity, this MONTHLY, 73 (1966) 969-973.
4. R. Sivaramakrishnan, Problem E 2196, this MONTHLY, 77 (1970) 772.

PERFECT PARALLELOGRAMS

R. W. SIELAFF, Naperville, Illinois

M. V. Subbarao [1] showed that the number of triangles (a, b, c) , whose integer-valued sides a, b, c add up to λ times their area, is finite for all positive λ . In fact he showed that with the exception of the triangle $(2, 2, 2)$ the number of such triangles (called Perfect Triangles) is zero for $\lambda > \sqrt{8}$. [*Editor's note:* In the latter part of Subbarao's proof, there is a mistake which has caused the triangles $(1, b, b)$ to be dropped from contention. For each of these, $\lambda > \sqrt{8}$.] He also suggested that it would be interesting to consider a similar problem for a quadrilateral. This note will consider a similar problem for a parallelogram.

Let a, b, c be positive integers greater than zero. If D is the area of a parallelogram with adjacent sides b and c and included angle A , then $D = bc \sin A$.

DEFINITION. A Perfect Parallelogram (b, c) is a parallelogram such that $2b + 2c = aD$. From the above, $abc \sin A = 2b + 2c$ or

$$\sin A = \frac{2}{a} \left(\frac{1}{b} + \frac{1}{c} \right).$$

Since $\sin A \leq 1$, $1/b + 1/c \leq a/2$.

The following solutions are possible:

$$a = 1, b = 3, c \geq 6; \quad a = 1, b = 4, c \geq 4$$

$$a = 1, b = 5, c \geq 4; \quad a = 1, b \geq 6, c \geq 3$$

$$a = 2, b \geq 2, c \geq 2$$

$$a = 3, b = 1, c \geq 2; \quad a = 3, b \geq 2, c = 1$$

$$a \geq 4, b \geq 1, c \geq 1.$$

From the above discussion it is clear that there are infinitely many perfect parallelograms.

For the special case of the rectangle, $\sin A = 1$ and $c = b + k$, where k is an integer, $k \geq 0$:

$$a = \frac{2}{b} + \frac{2}{b+k}.$$

For $b > 4$ there are no perfect rectangles since $a < 1$ for any k . For $b = 1, 2, 3, 4$, there are five perfect rectangles:

$$(1,1), (1,2), (2,2), (3,6), (4,4).$$

For the special case, $\sin A = \frac{1}{2}$ and $c = b + k$

$$a = \frac{4}{b} + \frac{4}{b+k}.$$

For $b > 8$ there are no such perfect parallelograms since $a < 1$ for any k . For $b = 7$ there is no solution. For $b = 1, 2, 3, 4, 5, 6, 8$, there are ten such perfect parallelograms:

$$(1,1), (1,2), (1,4), (2,2), (2,4), (3,6), (4,4), (5,20), (6,12), (8,8).$$

I would like to acknowledge with thanks the suggestions of the referee.

Reference

1. M. V. Subbarao, Perfect Triangles, this MONTHLY, 78 (1971) 384-385.

A CROWDED SET OF NON-INTERSECTING LINES

J. A. EIDSWICK, University of Nebraska

THEOREM. *There exists a family of lines, uncountably many in each of an uncountable number of directions, which has no intersection points on the strip*

$$S = \{(x, y): 0 \leq x \leq 1, -\infty < y < \infty\}.$$

Proof. Define h on the Cantor ternary set C by

$$h\left(\sum_{n=1}^{\infty} a_n 3^{-n}\right) = \sum_{n=1}^{\infty} a_{2n} 9^{-n}$$

and extend h linearly to all of the interval $[0, 1]$. Obviously, there are uncountably many values each of which is attained by h uncountably many times. Also it is easy to show that the function $h(t) + t$ is increasing on C and therefore on $[0, 1]$. (For additional properties of h and the existence of smoother "uncountably recurrent functions" see [1].)

Now for each $t \in [0, 1]$, let $L(t)$ be the line defined by $y = h(t)x + t$, and let $\mathcal{F} = \{L(t): 0 \leq t \leq 1\}$.

Reference

1. R. B. Darst, C^∞ -functions need not be bimeasurable, Proc. Amer. Math. Soc., 27 (1971) 128-132.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

A DECEPTION GAME

JOEL SPENCER, University of California, Los Angeles

The following game, and problem, are due to Mark Thompson. The problem has been worked on by a number of mathematicians with little success.

The Deception Game is a two person zero-sum game. The first move is a chance move. An ordered triple of numbers x_1, x_2, x_3 is chosen independently from a uniform distribution on $[0, 1]$. The second move is the "deception move". Player II looks at x_1, x_2, x_3 and changes one of the x 's to some number (possibly the same). More formally, he picks y_1, y_2, y_3 so that $x_i = y_i$ for at least two of the values $i = 1, 2, 3$. The third, and final, move is the "guessing move". Player I looks at y_1, y_2, y_3 and picks $i = 1, 2$, or 3 . The payoff to Player I is x_i .

Let w be the value of the game (if it exists). Player I may assure himself $\frac{1}{2}$ by the strategy "Choose $i = 1$ ". (Or any other strategy that does not involve looking at y_1, y_2, y_3 .) Thus $w \geq 1/2$.

BIG QUESTION: Does $w = 1/2$?

That is, can Player II so bamboozle his opponent that Player I can do no better than guess without looking at y_1, y_2, y_3 ?

If $w = 1/2$ there is a strategy for Player II that holds Player I to $1/2$. (Or possibly, for $\varepsilon > 0$, a strategy that holds Player I to $1/2 + \varepsilon$.) A necessary and sufficient condition on such a strategy is that when Player I sees y_1, y_2, y_3 he must have preference. That is $E[x_i | y_1, y_2, y_3]$ is independent of i .

We note that the computation of the payoff given strategies for Players I and II often involves subtle probabilistic considerations.

The reader might find it instructive to show that the following strategies for Player II do not hold Player I to $w = 1/2$. We assume the x 's satisfy $x_{i_1} < x_{i_2} < x_{i_3}$.

- (1) Change x_{i_3} to x_{i_1}
- (2) Change x_{i_3} to y_{i_3} picked from uniform distribution on $[0, 1]$
- (3) Change x_{i_2} to y_{i_2} picked from uniform distribution on $[0, x_{i_1}] \cup [x_{i_3}, 1]$.

As this is an infinite game it is not clear, and has not been proved, that a value w exists. The usual compactness arguments on the strategy space do not seem to work although one certainly feels, intuitively, that w does exist.

Known results. (All results on this problem are unpublished. The results of Mark Thompson are part of his unpublished undergraduate thesis—Harvard University, 1970.)

Thompson considered a discrete version when the x_i are chosen uniformly from

$$\left\{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n}{n}\right\} \text{ and proved } w = 1/2 \text{ for } 1 \leq n < 5.$$

It seems reasonable to generalize and allow x_1, x_2, x_3 to come from any (but the same) distribution of finite mean. Then the conjecture would be that the value w is that mean. This writer, for example, showed the conjecture true if the x_i are chosen uniformly from $\{0, 1, 8\}$.

Some work has been done on the generalization where n numbers x_1, \dots, x_n are chosen uniformly in $[0, 1]$ and k of them are changed by Player II. E. B. Keeler proved that if $n \leq 2k$, $w = 1/2$. D. Kleitman and, independently, S. Zamir, proved that if $n = 4$, $k = 1$ then $w > 1/2$.

CLASSROOM NOTES

EDITED BY ROBERT GILMER

Manuscripts for this Department should be sent to Robert Gilmer, Department of Mathematics, Florida State University, Tallahassee, FL 32306. Notes are usually limited to three printed pages.

TRAFFIC FLOW: LAPLACE TRANSFORMS

E. A. BENDER AND L. P. NEUWIRTH, Institute for Defense Analyses

Preface. We have found that students seem to develop a better appreciation of differential equations when presented with substantial applications. The following project is a modification of a handout we have used in our classes. Each question is preceded by "Q". Selected answers appear at the end. Section IV is written at an elementary, heuristic level, because our classes were studying differential equations as part of the basic calculus sequence. The handout referred to in VI applied stability theory to the Volterra-Lotka equations for predator-prey models and to a pendulum with a frictional force which is an arbitrary function of velocity.

1. Introduction. The mathematical study of traffic flow is relatively new. It requires little scientific background and uses primarily differential equations, probability and statistics. A book [1] is available; however, you may find it difficult if you have not had some probability and statistics. There are also brief surveys in [2, 3].

We shall consider the following problem:

A single line of cars moves along a straight highway without passing. Under what conditions will acceleration and/or deceleration by the first driver cause a collision further back?

It is clear that *very rapid* action by the first driver can easily cause a pile up. We are interested in situations in which moderate action by the first driver causes more and more violent responses as the effect travels back along the line of cars.

2. The model. This model is taken from [4]. The position of the lead car is given by $x_1(t)$, the position of the n th by $x_n(t)$. All drivers will be treated as identical. (This is not essential. It only simplifies calculations.) Time is measured in units of driver reaction time. Each driver's acceleration (deceleration) is proportional to the difference between the speed of his car and that of the car ahead. The first driver is free to do as he wishes. Thus we assume

$$(1) \quad x_n''(t) = C(x_{n-1}'(t-1) - x_n'(t-1))$$

for some $C > 0$ and all $n > 1$. The $t-1$ is due to reaction time lag. For $t \leq 0$ we assume that $x_n'(t)$ is a constant independent of n ; that is, the string of cars is moving, as a unit with constant velocity.

Q1: Comment on the reasonableness of (1). In particular, might C depend on car separation? In a qualitative way, how? Might it differ for acceleration and deceleration? How?

Q2: Introduce $z_n(t) = x_n(t) - x_n(0) - tx_n'(0)$. Interpret z_n and show that

$$(2) \quad \begin{cases} z_n''(t) = C(z_{n-1}'(t-1) - z_n'(t-1)) & \text{for } n > 1, \\ z_n(t) = z_n'(t) = 0 & \text{for } t \leq 0 \text{ and } n \geq 1. \end{cases}$$

(The fact that $z_n(0) = z_n'(0) = 0$ simplifies the next step.)

3. Some Laplace transforms.

Q3: Assume that the lead car varies its speed in some fashion when $t > 0$. Take Laplace transforms and show that

$$(3) \quad Z_{n+1}(s) = C^n(C + se^s)^{-n} Z_1(s),$$

where $Z_n(s) = \mathcal{L}(z_n(t))$.

Q4: Let $a_n(t) = z_n''(t) = x_n''(t)$, the acceleration of the n th car. Denote $\mathcal{L}(a_n(t))$ by $A_n(s)$. Using (3) deduce

$$(4) \quad A_{n+1}(s) = C^n(C + se^s)^{-n} A_1(s).$$

(You should recognize that this formula expresses the Laplace transformed description of the $n+1$ car's behaviour in terms of the first car's behaviour!) When $A_1(s)$ is specified, these transforms can be inverted, (with work), but this

approach leads to something messy and hard to use. Instead, one may rely on a bit of the theory of complex variables to obtain some simple approximate results.

4. Approximate inversion of Laplace transforms. We want to know roughly how the inverse transform of (4) behaves. In this section, we discuss without proof a well-known fact in the theory of Laplace transforms.

The transform $\mathcal{L}(e^{at}) = (s - a)^{-1}$ was derived in class for s a real number greater than a . However, $(s - a)^{-1}$ makes sense for all complex numbers $s \neq a$. All Laplace transforms can be extended to complex numbers in this fashion.

Let $g(s) = 1/\mathcal{L}(f)$. Suppose $g(s_0) = 0$ and no solution of $g(s) = 0$ has larger real part. Also suppose

$$g'(s_0) = g''(s_0) = \cdots = g^{(K)}(s_0) = 0$$

and $g^{(K+1)}(s_0) \neq 0$. One calls s_0 a zero of g of multiplicity $K + 1$. Let $s_0 = a + bi$.

FACT: Under the above assumptions there is a constant E such that $f(t)$ grows like $E t^K e^{at}$. In addition, $f(t)$ may oscillate. If $b \neq 0$, then there is oscillation like $\cos(bt + d)$ for some d . Of course, there may be several roots of $g(s) = 0$ with real part a and different imaginary parts. Then there will be several components of the oscillation.

Examples. (The facts given here are without proof, you may wish to verify them.) Suppose

$$\mathcal{L}(f) = \left(\frac{1}{(s-a)^2 + b^2} \right) \left(\frac{1}{(s-a)^2 + b'^2} \right)$$

$b, b' \neq 0$. Then $g(s) = ((s-a)^2 + b^2)((s-a)^2 + b'^2)$ has roots $a \pm bi, a \pm b'i$. At each root $g'(s) \neq 0$. Hence $f(t)$ grows like

$$ce^{at} \cos(bt + d) + c'e^{at} \cos(b't + d')$$

for some constants c, c', d, d' . You can check this by finding $f(t)$. If $b' = 0$, then we get

$$ce^{at} \cos(bt + d) + \boxed{c'te^{at}}$$

and the term in the box is the important one. If $g(s) = ((s-a)^2 + b^2)(s-a)$, then $f(t)$ grows like

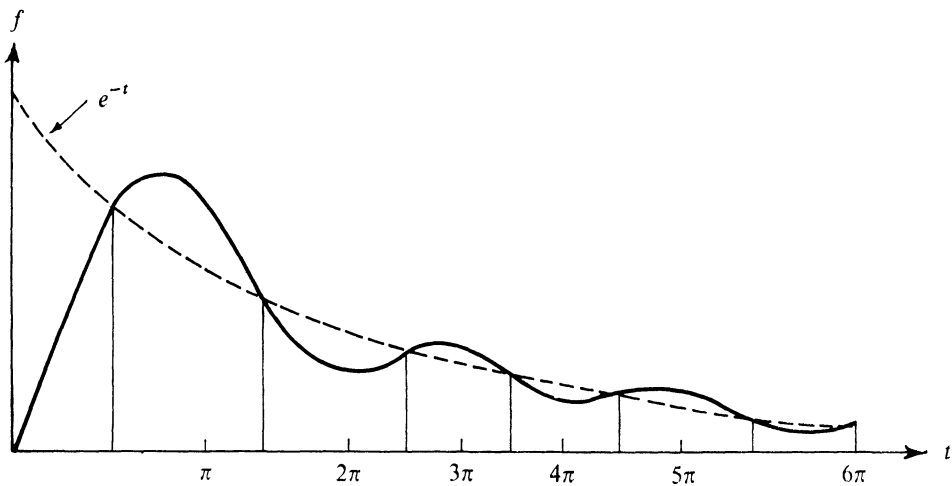
$$ce^{at} \cos(bt + d) + c'e^{at}.$$

Q5: Suppose $f = e^{-t} - e^{-2t} \cos t$. Then (complete this equation)

$$g(s) = \frac{1}{\mathcal{L}(f)} =$$

and $g(s) = 0$ has solutions $s =$

In this case, there is some oscillation due to the term $e^{-2t}\cos t$, but it dies out faster than e^{-t} . The graph of f is like that shown below. This relative unimportance of the oscillations is typical of the case in which there is only one root with largest real part and this root is a real number.



5. Analysis of our model using the previous sections. Now suppose the first driver accelerates and decelerates for a finite period of time; that is, $a_1(t) > 0$ for a bit, < 0 for a bit, and then $a_1(t) = 0$ for $t \geq T$ for some T .

Q6: Use the physical behaviour of the first car to show that $1/(A_1(s)) = 0$ has no solutions.

The following can be shown concerning the zeros of $C + se^s$ which have largest real part equal to a .

Condition	Nature of $a + bi$	
$C > \pi/2$	$a > 0$	$b \neq 0$
$C = \pi/2$	$a = 0$	$b \neq 0$
$\pi/2 > C$	$a < 0$	$b \neq 0$
$1/e \geq C$	$a < 0$	$b = 0$

Q7: Using the ideas developed in this handout, analyze each of the four cases for C . Describe the nature of $a_n(t)$, ($n > 1$) for large t , where the first driver behaves as suggested at the beginning of this section. Show that a collision *must* take place if $C \geq \pi/2$.

6. Some remarks. The above system is a “feedback mechanism” which can be characterized as follows:

- (i) a critical point (where $z_n(t) = 0$ for all n),
- (ii) a “mechanism” which acts to restore the system when it deviates from the critical point (here the equations (2)),
- (iii) often a time delay (here reaction time).

Similar ideas were studied in the stability theory handout where we studied the nature of critical points for non-linear models by linearization. A non-linear time delay situation can also be handled by a linearization approach, but it requires Laplace transforms as used here. Related ideas are presented in [6] and the references given there.

Feedback delay is important for stability, as the problem below shows. Suppose we measure things in units of ordinary time with Δ the reaction time. We have

$$(5) \quad x''_{n+1}(t) = k(x'_n(t-\Delta) - x'_n(t-\Delta)).$$

- Q8: (a) Transform (5) to the form (1) and so express C in terms of k and Δ .
 (b) Deduce that as reaction time increases the system tends to become unstable.
 (c) Now assume $\Delta = 0$. Thus replace $t - 1$ by t in (1). Show that

$$Z_{n+1}(s) = C^n(C + s)^{-n} Z_1(s).$$

Deduce that the system is stable regardless of the value of $C > 0$. (In fact, large C gives greater stability in direct opposition to the case $\Delta \neq 0$!).

Another sort of instability can occur in our example. A plot of $z_n(t)$ versus time may level off as $t \rightarrow \infty$, but as n gets larger, the graphs may become wilder. This can be studied in various ways:

- (i) Choose $x_1(t)$ so that (1) can be solved [5].
- (ii) Study the inverse transform of (4) by approximate methods which are more accurate than those used above [4, 5].
- (iii) Use other transforms to study other functions besides $a_n(t)$, for example $\int_0^\infty a_n(t)^2 dt$ [4].

We shall consider (i).

Let $x_n(t) = b_n e^{i\omega t}$ where b_n is to be found. In the end, we can let $x_n(t)$ be the real part of the above since $d/dt \operatorname{Re}(f(t)) = \operatorname{Re}(f'(t))$. (Re denotes “real part of.”) This gives a sinusoidal motion.

Q9: (a) Let $f(t)$ be an arbitrary differentiable complex valued function of the real number t . Show that $d/dt \operatorname{Re} f(t) = \operatorname{Re}(f'(t))$.

- (b) Using (1), show that

$$-\omega^2 b_n = iC\omega e^{-i\omega} (b_{n-1} - b_n).$$

- (c) Deduce that

$$b_{n+1} = b_1 / \left(1 + \frac{i\omega}{C} e^{i\omega} \right)^n.$$

(d) Show that we have instability if and only if

$$\left| 1 + \frac{i\omega}{C} e^{i\omega} \right| < 1$$

and hence when $C > \omega/2 \sin \omega$.

(e) Conclude that if $C > \frac{1}{2}$, there is instability for ω near zero.

Note the difference between the result $C > \frac{1}{2}$ and our earlier result $C > \pi/2 = 1.57 \dots$.

Q10: Account for the difference just noted.

Experiments indicate that C is nearly $\frac{1}{2}$ for the actual drivers. See [4]. We have only considered one type of motion for the leader, but the ideas can be generalized by using Fourier series or approach (ii). Approach (iii) has another advantage. We can allow for the fact that (1) should be replaced by a less deterministic equation. This involves some elementary statistics and the Fourier transform.

SELECTED ANSWERS

A4: We have

$$\mathcal{L}(z_n''(t)) = C\{\mathcal{L}(z_{n-1}'(t-1)) - \mathcal{L}(z_n'(t-1))\}.$$

By tables of \mathcal{L}

$$s^2 Z_n(s) = C(e^{-s} s Z_{n-1}(s) - e^{-s} s Z_n(s)).$$

Hence $Z_n(s) = C(C + e^s)^{-1} Z_{n-1}(s)$.

A6: Let $\alpha = \max_{0 \leq t \leq T} |a_1(t)|$ and $\beta(s) = \max_{0 \leq t \leq T} |e^{st}|$. Then

$$\begin{aligned} |A_1(s)| &\leq \int_0^\infty |e^{st} a_1(t)| dt = \int_0^T |e^{st} a_1(t)| dt \\ &\leq \int_0^T \beta(s) \alpha dt = \alpha \beta(s) T. \end{aligned}$$

Hence $(A_1(s))^{-1}$ has no roots.

A7: When $C > \pi/2$ acceleration is unstable: it oscillates with larger and larger amplitudes as time goes on. When $C = \pi/2$, there is stable oscillatory behavior for the second car and unstable oscillatory behavior for the rest of the string since

$$\frac{d}{ds} \frac{(C + se^s)^n}{A_1(s)} \bigg|_{s=a+bi} = 0 \text{ for } n > 1.$$

A8: (a) Let the independent variable be $\tau = t/\Delta$. We obtain

$$x_{n+1}''(\tau) = \Delta k(x_n'(\tau) - x_n'(\tau-1))$$

so the preceding analysis applies with $C = \Delta k$.

A10: We have considered *different* accelerations for the first car in the two cases. We have also given *different* answers.

- (a) In Section 5 we saw that the third car always develops wild acceleration if $C > \pi/2$.
- (b) By Q9(c) we see that no car *ever* develops wild acceleration, but each car is wilder than its predecessor if $C > \omega/2\sin\omega$.

References

1. F. Haight, *Mathematical Theories of Traffic Flow*, Academic Press, New York, 1963.
2. D. Gazis, *Mathematical theory of automobile traffic*, *Science*, 157 (7-21-67) 273-281.
3. R. Herman and Gardels, *Vehicular traffic flow*, *Scientific American*, 209 no. 6 (Dec. 1963) 35-43.
4. R. Herman, Montroll, Potts, and Rothery, *Traffic dynamics: Analysis of stability in car following*, *Operations Res.*, 7 (1959) 86-103.
5. R. Chandler, Herman, and Montroll, *Traffic Dynamics: Studies in car following*, *Operations Res.*, 6 (1958) 165-184.
6. H. Simon, *Models of Man*, Chapter 13, *Application of Servomechanism Theory to Production Control*, Wiley, New York, 1957.

IRRATIONAL NUMBERS

J. P. JONES AND S. TOPOROWSKI, University of Calgary

For the past few years a clever proof has been making the rounds of the various mathematics departments.

THEOREM 1. *An irrational number raised to an irrational power may be rational.*

Proof: Consider the identity

$$[\sqrt{2^{\sqrt{2}}}]^{\sqrt{2}} = 2.$$

If $\sqrt{2^{\sqrt{2}}}$ is rational then we are finished. If not then $\sqrt{2^{\sqrt{2}}}$ is irrational so $(\sqrt{2^{\sqrt{2}}})^{\sqrt{2}}$ is the example.

This proof seems first to have been published by Dov Jarden as a curiosity in [3]. The proof was published again in [2]. Note that while the proof is elementary, it is non-constructive. The non-constructivity enters in the form of the logical principle of the excluded middle (*tertium non datur*) which the intuitionists reject.

Actually $\sqrt{2^{\sqrt{2}}}$ is irrational, being the square root of Hilbert's number $2^{\sqrt{2}}$, proved transcendental by Kuzmin [1] in 1930. But this result, which is not elementary, is not used above. Only the irrationality of $\sqrt{2}$ is used.

Consider next the related theorem.

THEOREM 2. *An irrational number raised to an irrational power may be irrational.*

Of course we can use set theoretical principles to prove that a^b is irrational for almost all real numbers b . Or we can use the result of Kuzmin [1] to prove Theorem 2. But does Theorem 2 have an elementary proof?

Proof: Consider the identity $\sqrt{2^{(\sqrt{2}+1)}} = (\sqrt{2^{\sqrt{2}}}) \sqrt{2}$.

If $\sqrt{2^{\sqrt{2}}}$ is irrational then we are finished. If not, then $\sqrt{2^{\sqrt{2}}}$ is rational. Hence $(\sqrt{2^{\sqrt{2}}})\sqrt{2}$ is irrational, and $\sqrt{2^{(\sqrt{2}+1)}}$ is the example in this case.

There is also a simple identity by means of which it can be proved that a rational number raised to an irrational power may be irrational. But perhaps the reader would enjoy finding this one himself.

References

1. R. Kuzmin, On a new class of transcendental numbers, *Izv. Akad. Nauk SSSR, Ser. Mat.*, 7 (1930) 585–597.
2. *Mathematics Magazine*, 39(1966) 111, 134.
3. *Scripta Mathematica*, 19 (1953) 229.

$$\text{A SIMPLE PROOF OF THE FORMULA } \sum_{k=1}^{\infty} k^{-2} = \pi^2/6$$

IOANNIS PAPADIMITRIOU, Athens, Greece

Start with the inequality $\sin x < x < \tan x$ for $0 < x < \pi/2$, take reciprocals, and square each member to obtain

$$\cot^2 x < 1/x^2 < 1 + \cot^2 x.$$

Now put $x = k\pi/(2m+1)$ where k and m are integers, $1 \leq k \leq m$, and sum on k to obtain

$$(1) \quad \sum_{k=1}^m \cot^2 \frac{k\pi}{2m+1} < \frac{(2m+1)^2}{\pi^2} \sum_{k=1}^m \frac{1}{k^2} < m + \sum_{k=1}^m \cot^2 \frac{k\pi}{2m+1}.$$

But since we have

$$(2) \quad \sum_{k=1}^m \cot^2 \frac{k\pi}{2m+1} = \frac{m(2m-1)}{3},$$

(a proof of (2) is given below) relation (1) gives us

$$\frac{m(2m-1)}{3} < \frac{(2m+1)^2}{\pi^2} \sum_{k=1}^m \frac{1}{k^2} < m + \frac{m(2m-1)}{3}.$$

Multiply this relation by $\pi^2/(4m^2)$ and let $m \rightarrow \infty$ to obtain

$$\lim_{m \rightarrow \infty} \sum_{k=1}^m \frac{1}{k^2} = \frac{\pi^2}{6}.$$

Proof of (2). By equating imaginary parts in the formula

$$\cos n\theta + i \sin n\theta = (\cos \theta + i \sin \theta)^n = \sin^n \theta (\cot \theta + i)^n$$

$$= \sin^n \theta \sum_{k=0}^n \binom{n}{k} i^k \cot^{n-k} \theta,$$

we obtain the trigonometric identity

$$\sin n\theta = \sin^n \theta \left\{ \binom{n}{1} \cot^{n-1} \theta - \binom{n}{3} \cot^{n-3} \theta + \binom{n}{5} \cot^{n-5} \theta - + \dots \right\}.$$

Take $n = 2m + 1$ and write this in the form

$$(3) \quad \sin(2m+1)\theta = \sin^{2m+1} \theta P_m(\cot^2 \theta) \text{ with } 0 < \theta < \frac{\pi}{2},$$

where P_m is the polynomial of degree m given by

$$P_m(x) = \binom{2m+1}{1} x^m - \binom{2m+1}{3} x^{m-1} + \binom{2m+1}{5} x^{m-2} - + \dots.$$

Since $\sin \theta \neq 0$ for $0 < \theta < \pi/2$, equation (3) shows that $P_m(\cot^2 \theta) = 0$ if and only if $(2m+1)\theta = k\pi$ for some integer k . Therefore $P_m(x)$ vanishes at the m distinct points $x_k = \cot^2 \pi k / (2m+1)$ for $k = 1, 2, \dots, m$. These are all the zeros of $P_m(x)$ and their sum is

$$\sum_{k=1}^m \cot^2 \frac{\pi k}{2m+1} = \binom{2m+1}{3} / \binom{2m+1}{1} = \frac{m(2m-1)}{3},$$

which proves (2).

NOTE. This paper was translated from a Greek manuscript and communicated to the MONTHLY on behalf of the author by Tom M. Apostol, California Institute of Technology. After this paper was written it was learned that the same proof was discovered independently and published in Norwegian by Finn Holme in *Nordisk Matematisk Tidskrift*, vol. 18 (1970), pp. 91–92. See also A. M. Yaglom and I. M. Yaglom, *Challenging mathematical problems with elementary solutions*, vol. II, Holden-Day, San Francisco, 1967, problem 145.

ANOTHER ELEMENTARY PROOF OF EULER'S FORMULA FOR $\zeta(2n)$

TOM M. APOSTOL, California Institute of Technology

1. Introduction. The classic formula

$$(1) \quad \zeta(2n) = \sum_{k=1}^{\infty} \frac{1}{k^{2n}} = (-1)^{n-1} \frac{(2\pi)^{2n} B_{2n}}{2(2n)!}$$

which expresses $\zeta(2n)$ as a rational multiple of π^{2n} was discovered by Euler [2]. The numbers B_n are Bernoulli numbers and can be defined by the recursion formula

$$B_0 = 1, \quad B_n = -\sum_{s=0}^n \binom{n}{s} B_s \text{ for } n \geq 2,$$

or equivalently, as the coefficients in the power series expansion

$$(2) \quad \frac{z}{e^z - 1} = \sum_{n=0}^{\infty} \frac{B_n z^n}{n!}, \quad |z| < 2\pi.$$

In this notation we have

$$(3) \quad B_1 = -\frac{1}{2}, \quad B_{2n+1} = 0 \text{ for } n \geq 1,$$

and

$$B_2 = \frac{1}{6}, \quad B_4 = -\frac{1}{30}, \quad B_6 = \frac{1}{42}, \quad B_8 = -\frac{1}{30}, \quad B_{10} = \frac{5}{66}.$$

Euler's original proof of (1) was obtained from two distinct representations of $\pi z \cot \pi z$, a power series expansion obtainable from (2),

$$\pi z \cot \pi z = 1 + \sum_{n=1}^{\infty} (-1)^n \frac{(2\pi z)^{2n} B_{2n}}{(2n)!}, \quad \text{valid for } |z| < 1,$$

and the partial fraction decomposition

$$\pi z \cot \pi z = 1 - 2 \sum_{k=1}^{\infty} \frac{z^2}{k^2 - z^2}, \quad \text{valid for } z \neq 0, \pm 1, \pm 2, \dots$$

If $|z| < 1$, each term in the last sum can be expanded in a geometric series giving us

$$\pi z \cot \pi z = 1 - 2 \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} \left(\frac{z^2}{k^2} \right)^n = 1 - 2 \sum_{n=1}^{\infty} \zeta(2n) z^{2n}.$$

Equation (1) follows by equating coefficients of z^{2n} in the two power series expansions of $\pi z \cot \pi z$. Details justifying this argument are given in Knopp [4], pp. 203–207, 236.

Another well-known proof is obtained by putting $s = 2n$ in Riemann's functional equation

$$\zeta(1-s) = 2(2\pi)^{-s} \Gamma(s) \cos \frac{\pi s}{2} \zeta(s)$$

and using the fact that $\zeta(1-2n) = -B_{2n}/(2n)$. These results are deduced by applying residue calculus to a contour integral representation of $\zeta(s)$. (See Titchmarsh [8], pp. 18–20.)

Several writers have given more elementary proofs of (1) that do not require concepts from advanced real or complex analysis. For example, Titchmarsh [7] obtained a set of complicated recursion formulas which can be used to evaluate $\zeta(4), \zeta(6), \dots$, successively in terms of $\zeta(2)$. Estermann [1] obtained a simpler formula of the same type. These recursion formulas, which show that $\zeta(2n)$ is a rational multiple of $\zeta(2)^n$, were deduced by rearranging absolutely convergent infinite series but did not require any function theory. Estermann also gave an elementary proof of the formula $\zeta(2) = \pi^2/6$ as a consequence of Gregory's series $1 - \frac{1}{3} + \frac{1}{5} - \dots = \pi/4$.

A recursion formula simpler than those of Titchmarsh and Estermann was proved by G. T. Williams [11] who showed by elementary methods that

$$(4) \quad \left(n + \frac{1}{2}\right) \zeta(2n) = \sum_{k=1}^{n-1} \zeta(2k) \zeta(2n-2k).$$

He also obtained the companion result

$$(5) \quad \left(n - \frac{1}{2}\right) (1 - 2^{-2n}) \zeta(2n) = \sum_{k=1}^n \zeta(2k-1) \zeta(2n-2k+1),$$

where

$$\xi(s) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)^s} \text{ for } s > 0.$$

Note that $\xi(1)$ is Gregory's series for $\pi/4$. Taking $n = 1$ in (5) we find $\frac{3}{8}\zeta(2) = \xi^2(1) = \pi^2/16$, so $\zeta(2) = \pi^2/6$. This result, in conjunction with (4), gives a completely elementary evaluation of $\zeta(2n)$ as a rational multiple of π^{2n} . Williams also points out that (4) is equivalent to the following recursion formula for Bernoulli numbers,

$$-(2n+1)B_{2n} = \sum_{k=1}^{n-1} \binom{2n}{2k} B_{2k} B_{2n-2k}.$$

This relation appears in Nielsen's book [5] and was also discovered independently by R. S. Underwood [9] who used it to evaluate the sums $\sum_{k=1}^m k^n$ in terms of Bernoulli polynomials.

The purpose of this note is to show that the elementary method used by Papadimitriou to evaluate $\zeta(2)$ in the foregoing paper [6] can be extended to evaluate $\zeta(2n)$ and leads directly to Equation (1) rather than to a recursion formula. The interplay of ideas from elementary algebra and trigonometry makes the proof especially suitable for an elementary calculus course.

2. Elementary Proof of (1). The key ingredient in Papadimitriou's proof is the formula

$$\sum_{k=1}^m \cot^2 \frac{k\pi}{2m+1} = \frac{m(2m-1)}{3},$$

or rather the asymptotic relation

$$(6) \quad \sum_{k=1}^m \cot^2 \frac{k\pi}{2m+1} = \frac{2}{3}m^2 + O(m)$$

which it implies. Our evaluation of $\zeta(2n)$ makes use of the following lemma which provides a generalization of (6).

LEMMA 1. *For any integers $m \geq 1, n \geq 1$, we have*

$$(7) \quad \sum_{k=1}^m \cot^{2n} \frac{k\pi}{2m+1} = (-1)^{n-1} \frac{2^{4n-1} B_{2n}}{(2n)!} m^{2n} + O(m^{2n-1}),$$

where the constant implied by the O -symbol is independent of m .

First we show how the lemma implies (1) and then we prove the lemma.

The inequality $\sin x < x < \tan x$ for $0 < x < \pi/2$ implies

$$\cot^{2n} x < \frac{1}{x^{2n}} < (1 + \cot^2 x)^n$$

for each integer $n \geq 1$. We take $x = k/(2m+1)$ and sum on k to obtain

$$(8) \quad \sum_{k=1}^m \cot^{2n} \frac{k\pi}{2m+1} < \frac{(2m+1)^{2n}}{\pi^{2n}} \sum_{k=1}^m \frac{1}{k^{2n}} < \sum_{k=1}^m \left(1 + \cot^2 \frac{k\pi}{2m+1}\right)^n.$$

From (7) and the binomial theorem we see that

$$\sum_{k=1}^m \left(1 + \cot^2 \frac{k\pi}{2m+1}\right)^n = \sum_{k=1}^m \cot^{2n} \frac{k\pi}{2m+1} + O(m^{2n-1}).$$

Therefore if we multiply (8) by $\pi^{2n}/(2m)^{2n}$ and let $m \rightarrow \infty$ we obtain

$$\lim_{m \rightarrow \infty} \sum_{k=1}^m \frac{1}{k^{2n}} = (-1)^{n-1} \frac{(2\pi)^{2n} B_{2n}}{2(2n)!},$$

which proves (1).

3. Proof of Lemma 1. As in Papadimitriou's paper we use the polynomial

$$P_m(x) = \binom{2m+1}{1} x^m - \binom{2m+1}{3} x^{m-1} + \binom{2m+1}{5} x^{m-2} - + \dots$$

whose m zeros are the numbers

$$x_k = \cot^2 \frac{k\pi}{2m+1}, \quad k = 1, 2, \dots, m.$$

Let $s_n = x_1^n + \dots + x_m^n$. This sum appears on the left of (7) and we are to prove that

$$(9) \quad s_n = (-1)^{n-1} \frac{2^{4n-1} B_{2n}}{(2n)!} m^{2n} + O(m^{2n-1}).$$

The proof is by induction on n . The case $n = 1$ was proved in Papadimitriou's paper. Now we assume that (9) is true for $n = 1, 2, \dots, r-1$, and prove it for $n = r$, with the help of Newton's formulas (see [10], p. 261)

$$(10) \quad -s_r = (-1)^r r \sigma_r + \sum_{k=1}^{r-1} (-1)^{r-k} s_k \sigma_{r-k}, \quad r = 1, 2, \dots, m,$$

where $\sigma_1, \sigma_2, \dots, \sigma_m$ are the elementary symmetric functions of the zeros x_1, \dots, x_m . In this case we have

$$(11) \quad \sigma_r = \binom{2m+1}{2r+1} / \binom{2m+1}{1} = \frac{2m(2m-1)\cdots(2m-2r+1)}{(2r+1)!} \\ = \frac{2^{2r}}{(2r+1)!} m^{2r} + O(m^{2r-1}),$$

for $r = 1, 2, \dots, m$. Using this with (9) we find

$$(-1)^{r-k} s_k \sigma_{r-k} = (-1)^{r-1} \frac{2^{2r+2k-1} B_{2k}}{(2k)!(2r+1-2k)!} m^{2r} + O(m^{2r-1}),$$

so (10) becomes

$$-s_r = \frac{2r(-1)^r 2^{2r-1}}{(2r+1)!} m^{2r} + (-1)^{r-1} 2^{2r-1} m^{2r} \sum_{k=1}^{r-1} \frac{2^{2k} B_{2k}}{(2k)!(2r+1-2k)!} + O(m^{2r-1}) \\ = (-1)^r 2^{2r-1} m^{2r} \left\{ \frac{1}{(2r)!} - \sum_{k=0}^{r-1} \frac{2^{2k} B_{2k}}{(2k)!(2r+1-2k)!} \right\} + O(m^{2r-1}).$$

Now we use Lemma 2 (stated below) to evaluate the expression in braces and we find

$$-s_r = (-1)^r \frac{2^{4r-1} B_{2r}}{(2r)!} m^{2r} + O(m^{2r-1}),$$

which proves (9) by induction.

4. A lemma on Bernoulli numbers.

LEMMA 2. If $r \geq 1$ we have

$$(12) \quad \sum_{k=0}^r \frac{2^{2k} B_{2k}}{(2k)!(2r+1-2k)!} = \frac{1}{(2r)!}.$$

Proof. Let $B_n(x)$ denote the Bernoulli polynomial defined by

$$(13) \quad B_n(x) = \sum_{s=0}^n \binom{n}{s} B_s x^{n-s}$$

or, equivalently, by the power series expansion

$$\frac{ze^{xz}}{e^z - 1} = \sum_{n=0}^{\infty} \frac{B_n(x)}{n!} z^n, \quad |z| < 2\pi.$$

A well-known property of $B_n(x)$ is the functional equation

$$(14) \quad B_n(1-x) = (-1)^n B_n(x)$$

which follows at once from the identity

$$\frac{ze^{(1-x)z}}{e^z - 1} = \frac{-ze^{-xz}}{e^{-z} - 1}.$$

Equation (14) implies

$$(15) \quad B_{2r+1} \left(\frac{1}{2} \right) = 0.$$

Formula (12) is a disguised form of (15). We use (15) along with (13) and multiply by 2^{2r+1} to obtain

$$\sum_{s=0}^{2r+1} \binom{2r+1}{s} 2^s B_s = 0.$$

In view of (3) this becomes

$$\binom{2r+1}{1} 2B_1 + \sum_{k=0}^r \binom{2r+1}{2k} 2^{2k} B_{2k} = 0,$$

which is the same as (12).

5. Concluding remarks. When the method of section 2 is applied to evaluate $\zeta(2n+1)$ we obtain the formula

$$(16) \quad \zeta(2n+1) = \left(\frac{\pi}{2} \right)^{2n+1} \lim_{m \rightarrow \infty} \frac{1}{m^{2n+1}} \sum_{k=1}^m \cot^{2n+1} \frac{k\pi}{2m+1},$$

or its equivalent,

$$(17) \quad \sum_{k=1}^m \cot^{2n+1} \frac{k\pi}{2m+1} = \left(\frac{2}{\pi} \right)^{2n+1} \zeta(2n+1) m^{2n+1} + o(m^{2n+1}) \text{ as } m \rightarrow \infty.$$

Although (16) expresses $\zeta(2n+1)$ as a multiple of π^{2n+1} it is not known if this multiple is rational or not. The author has been unable to extend the proof of Lemma 1 to obtain an alternate formula for the asymptotic value (for large m) of the sum in (17). All attempts to estimate this sum lead back to (17).

NOTE. After this paper was submitted for publication, a paper appeared by Kenneth S. Williams [12] on the same subject. Williams also uses the cotangent sum of Lemma 1 in his evaluation of $\zeta(2n)$, but his proof, like Euler's, uses complex function theory and cannot be considered elementary. See also I. Skau and E. S. Selmer, *Nordisk Mat. Tidskr.*, 19 (1971) 120-124.

References

1. T. Estermann, Elementary evaluation of $\zeta(2k)$, *J. London Math. Soc.*, 22 (1947) 10-13.
2. L. Euler, De summis serierum reciprocarum, *Comment. Acad. Sci. Petropolit.*, 7 (1734/35), (1740) 123-134; *Opera omnia*, Ser. 1 Bd. 14, 73-86. Leipzig-Berlin, 1924.
3. Finn Holme, En enkel beregning av $\sum_{k=1}^{\infty} 1/k^2$, *Nordisk Mat. Tidskr.*, 18 (1970) 91-92.
4. K. Knopp, *Theory and Application of Infinite Series*, Hafner, New York, 1951.
5. N. Nielsen, *Traité élémentaire des nombres de Bernoulli*, Gauthier-Villars, Paris, 1923.

6. Ioannis Papadimitriou, A simple proof of the formula $\sum_{k=1}^{\infty} k^{-2} = \pi^2/6$, this MONTHLY, 80 (1973) preceding article.
7. E. C. Titchmarsh, A series inversion formula, Proc. London Math. Soc., (2) 26 (1926) 1-11.
8. ———, The Theory of the Riemann Zeta Function, Oxford, 1951.
9. R. S. Underwood, An expression for the summation $\sum_{m=1}^n m^p$, this MONTHLY, 35 (1928) 424-428.
10. J. V. Uspensky, Theory of Equations, McGraw-Hill, New York, 1948.
11. G. T. Williams, A new method of evaluating $\zeta(2n)$, this MONTHLY, 60 (1953) 19-25.
12. Kenneth S. Williams, On $\sum_{n=1}^{\infty} (1/n^{2k})$, Math. Mag., 44 (1971) 273-276.

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to either of the editors: J. G. Harvey, Department of Mathematics, University of Wisconsin, WI 53706; M. W. Pownall, Department of Mathematics, Colgate University, Hamilton, NY 13346.

AN INTEGRATED SEQUENCE IN THE MATHEMATICAL SCIENCES FOR UNDERGRADUATE BUSINESS STUDENTS

R. H. RANGLES AND A. J. SCHAEFFER, University of Iowa

The courses in mathematical science (mathematics, statistics, and computer programming) which are required for every business student vary widely among colleges and universities. In a recent sample survey of midwestern universities, Rodger Collons [1] found that among the 30 schools surveyed on the semester system, the required hours fell between the extremes of 0 and 21. The median of the required hours among those 30 schools was 9. A typical program might therefore consist of one 3 hour course each in mathematics, statistics, and computer programming. It is the purpose of this article to describe a sequence of two 4 semester hour courses developed at the University of Iowa in which topics from the three areas of mathematics, statistics and computer programming are blended together in an effort to increase the motivation of each of these subject areas. It is hoped that in so doing, the student will acquire more of an overview of the mathematical sciences and how techniques from all three disciplines lend themselves (possibly in conjunction with one another) to the solution of business problems. This article contains some of the details of this sequence and some suggestions for integrating topics from the mathematical sciences.

1. **Course structure.** The number of students entering this sequence each year is

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before July 31, 1973. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed stamped postcards.

E 2408. *Proposed by Bernardo Recamán, Colegio San Carlos, Bogotá, Colombia*

A natural number is a *decimal Colombian number* if it cannot be written as $m + s(m)$ for any natural number m , where $s(m)$ denotes the sum of the digits of m when m is expressed in decimal notation. For example, 28 is not a decimal Colombian number since $28 = 23 + 2 + 3$, whereas 9 is a decimal Colombian number. Base n Colombian numbers are defined analogously.

Prove that in any base there are infinitely many Colombian numbers.

E 2409. *Proposed by A. V. Boyd, University of the Witwatersrand, Johannesburg, South Africa*

Sum the series

$$\sum_{n=0}^{\infty} \binom{2n}{n}^{-1} (4x)^n.$$

E 2410. *Proposed by Barry Wolk, University of Manitoba*

Evaluate

$$\sum_{r=0}^n \frac{(-1)^r}{(n+r)(2r+1)} \binom{n+r}{2r}.$$

E 2411. *Proposed by F. W. Barnes, University of Michigan*

Let G be a group. Give sufficient conditions on a and b so that $(xy)^a = x^a y^a$ and $(xy)^b = x^b y^b$ for all $x, y \in G$, force G to be commutative. The conditions must be general enough to imply the result for $a = 8$ and $b = 11$.

E 2412. *Proposed by Michael Goldberg, Washington, D.C.*

If a line segment AB of unit length is rotated 180° about the fixed end B to the position BA' , then the end A makes a track of length π . However, if the end B is allowed to move also, then the sum of the lengths of the tracks made by A and B can be shorter. Note that if a portion of the track is retraced, then the motion is increased, but the length of the track is not increased. What is the shortest total length of track needed to carry the line from AB to BA' ?

E 2413. *Proposed by C. B. Grosch, Control Data Corp., Minneapolis*

Consider an oblate spheroid and the circle in its equatorial plane which is the locus of the foci of meridian ellipses. Show that any ray that originates on the circle will be reflected to the circle after a reflection from the interior of the spheroid. (On the spheroid, the angle of incidence equals the angle of reflection.)

SOLUTIONS OF ELEMENTARY PROBLEMS

Fibonacci's Rabbits Run Again

E 2350 [1972, 393]. *Proposed by H. D. Ruderman, Hunter College High School*

A total of n fair coins are flipped and laid in a row. What is the probability that in the row neither the combination HTH nor the combination THT occurs anywhere?

I. *Solution by N. J. Fine, Pennsylvania State University.* Let A_n, B_n, C_n, D_n be the numbers of successful n -sequences ending with HH, HT, TH, TT respectively. Thus $A_2 = B_2 = C_2 = D_2 = 1$. The required probability $p_n = 2^{-n} (A_n + B_n + C_n + D_n)$. By adjoining an H to each sequence of type A_n or C_n we get one of type A_{n+1} , and every successful sequence of type A_{n+1} is obtained in this way. Hence

$$A_{n+1} = A_n + C_n.$$

Similarly, $B_{n+1} = A_n, C_{n+1} = D_n, D_{n+1} = D_n + B_n$. By symmetry, $D_n = A_n$ and $C_n = B_n$. Hence

$$A_{n+1} = A_n + B_n = A_n + A_{n-1} \quad (n \geq 3).$$

Let f_n be the n th Fibonacci number. Then $A_2 = 1 = f_2, A_3 = 2 = f_3$. Hence $A_n = f_n$ for all $n \geq 2$, and

$$p_n = 2^{-n}(2A_n + 2B_n) = 2^{-n+1}(A_n + A_{n-1}) = 2^{-n+1}A_{n+1},$$

so finally

$$p_n = \frac{f_{n+1}}{2^{n-1}}.$$

II. *Solution by R. J. Dickson, Lockheed Palo Alto Research Laboratory.* Mark an S or D between each pair of adjacent coins according as they show the same or different faces. A sequence of $n - 1$ Bernoulli trials with probability $\frac{1}{2}$ results, and the probability sought is the same as the probability that this sequence does not contain two successive D 's. This is a special case of the "problem of runs" treated by de Moivre in his "Doctrine of Chances"; cf. Uspensky, *Introduction to Mathematical Probability*, McGraw-Hill (1937), 77-84.

Also solved by the proposer and fifty-three others.

Other references for the S - D sequence of solution II include problem A-5 of the 1956 Putnam Exam, and problem E 2022 in this MONTHLY [1968, 1117] (located by D. M. Bloom), problem B-236

in the *Fibonacci Quarterly*, 10 (1972) 330 (by Graham Lord), and Ivan Niven, *Mathematics of Choice*, Random House (1965), 50–52 (by D. P. Sumner). J. V. Michalowicz presented a computer-generated list of the probabilities p_n which included $p_5 = 0.500$, $p_{13} = 0.092$, $p_{24} = 0.009$, $p_{35} = 0.0009$.

A Number-theoretic Inequality

E 2351 [1972, 394]. *Proposed by Stefan Porubsky, Comenius' University, Bratislava, Czechoslovakia*

Let ϕ denote Euler's totient function and let $\tau(n)$ denote the number of divisors of n . Show that

$$\phi(n) [\tau(n)]^2 \leq n^2$$

for all positive integers $n \neq 4$. For what n does equality hold?

Solution by M. G. Greening, University of New South Wales, Australia. Set $\phi(n) [\tau(n)]^2 / n^2 = f(n)$.

(i) Clearly $f(1) = 1$, $f(2) = 1$, $f(4) = 9/8$ and $f(3) = 8/9$.

(ii) $f(p) = 4(p-1)/p^2 < 1$ for p an odd prime.

Also $f(p^{\alpha+1})/f(p^\alpha) = p(\alpha+2)^2/(\alpha+1)^2 p^2 \leq 9/4p$ for $\alpha > 0$, whence induction shows that $f(p^\alpha) < 1$ for all positive α .

(iii) $f(2^\alpha) = (\alpha+1)^2/2^{\alpha+1} < 1$ for $\alpha > 3$. $f(8) = 1$.

(iv) From (i) and (iii), $f(4k) = 1$ demands $k = 2$ or $(k, 2) = 1$.

In the latter case, $f(k) = 1/f(4) = 8/9$, so that $3 \mid k$. As $f(3^\alpha) < 8/9$ for $\alpha > 1$, by (ii), and $f(p^\alpha) = 1$ demands $p = 2$, while f is multiplicative, the only solution of $f(k) = 8/9$ is $k = 3$.

Hence, equality holds only for $n = 1, 2, 8, 12$.

Also solved by the proposer and forty-one others.

A Monotone Decreasing Sequence

E 2352 [1972, 394]. *Proposed by Marlow Sholander, Case Western Reserve University*

For each positive integer n , define

$$Q_n = \left[1 + \frac{1}{n}\right]^{n^2} \frac{n!}{n^n \sqrt{n}}.$$

Show that the sequence $\{Q_n\}$ is monotonely decreasing and find its limit.

Solution by St. Olaf College Students. It is easily calculated that

$$\frac{Q_{n+1}}{Q_n} = \frac{\left(1 + \frac{1}{n+1}\right)^{(n+1)^2}}{\left(1 + \frac{1}{n}\right)^{n^2+n+1/2}} = \left(1 - \frac{1}{(n+1)^2}\right)^{(n+1)^2} \left(1 + \frac{1}{n}\right)^{n+1/2} = e^{S_1+S_2}$$

where

$$S_1 = - \sum_{k=1}^{\infty} \frac{1}{(k+1)(n+1)^{2k}}, \quad S_2 = \sum_{k=3}^{\infty} (-1)^k \left(\frac{1}{k} - \frac{1}{2(k-1)} \right) \frac{1}{n^{k-1}}$$

are familiar convergent series for all positive integers n .

Note that S_1 is a series of negative terms and S_2 is an alternating series whose terms in absolute value decrease monotonely. Consequently the sums of both series are less than their respective first terms. Thus it follows that

$$S_1 + S_2 < -\frac{1}{2(n+1)^2} + \frac{1}{12n^2} < 0, \quad n = 1, 2, \dots$$

This proves that $Q_{n+1}/Q_n < 1$; that is, $\{Q_n\}$ is monotonely decreasing.

Since $Q_n > 0$, it follows that the sequence $\{Q_n\}$ possesses a limit. The limit is easily calculated if $n!$ is replaced by the familiar $(n/e)^n \sqrt{2\pi n}$. Thus

$$\begin{aligned} \lim_{n \rightarrow \infty} Q_n &= \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^{n^2} e^{-n} \sqrt{2\pi} \\ &= \sqrt{2\pi} \lim_{n \rightarrow \infty} \exp\left(n^2 \ln\left(1 + \frac{1}{n}\right) - n\right) \\ &= \sqrt{2\pi} \exp\left(\lim_{n \rightarrow \infty} \left(-\frac{1}{2} + \frac{1}{3n} - \frac{1}{4n^2} + \dots\right)\right) \\ &= \sqrt{2\pi/e}. \end{aligned}$$

Also solved by M. T. Bird, Peter Bundschuh (Germany), Frederick Carty, R. J. Dickson, Michael Goldberg, Richard Groenewald, Emil Grosswald, Sidney Heller, G. A. Heuer, Hans Kappus (Switzerland), Barbara A. Keller, Charlotte Krauthamer (Austria), O. P. Lossers (Netherlands), B. E. Rhoades, G. S. Rogers, T. Salát (Czechoslovakia), Leroy Sathre, F. G. Schmitt, Jr., W. C. Sisarcick, F. C. Smith, T. A. R. Stettler (Switzerland), L. M. Young, and the proposer.

Optimal Sequence of Products

E 2353 [1972, 394]. *Proposed by J. G. Rau, Litton Systems, Culver City, California*

Given two sequences $\{a_1, a_2, \dots, a_n\}$ and $\{b_1, b_2, \dots, b_n\}$ of positive real numbers, find the permutation (j_1, \dots, j_n) of the integers $1, 2, \dots, n$ for which

$$\sum_{m=1}^n \sum_{k=1}^m b_{j_m} a_{j_k}$$

is a minimum.

I. Solution by W. O. J. Moser, University College, London, England. Let S denote the sum when

$$(j_1, j_2, \dots, j_n) = (1, 2, \dots, n);$$

let S' denote the sum when

$$(j_1, j_2, \dots, j_n) = (1, 2, \dots, k-1, k+1, k, k+2, k+3, \dots, n).$$

Then

$$S' = S + b_k a_{k+1} - b_{k+1} a_k = S + b_k b_{k+1} \left(\frac{a_{k+1}}{b_{k+1}} - \frac{a_k}{b_k} \right),$$

so $S' < S$ if and only if $a_{k+1}/b_{k+1} < a_k/b_k$.

Letting $r_i = a_i/b_i$, $i = 1, \dots, n$, we see that the sum is minimum when (j_1, j_2, \dots, j_n) is such that

$$r_{j_1} \leq r_{j_2} \leq \dots \leq r_{j_n}.$$

II. *Solution by R. J. Dickson, Lockheed Palo Alto Research Laboratory.* Set $c_{km} = a_{j_k} b_{j_m}$, $C = \|c_{km}\|$, and let $C = L + U + D$ denote the decomposition of C into lower, upper, and diagonal parts. Let $s(A)$ denote the (linear, homogeneous) functional equal to the sum of the elements of a matrix A . The sum to be minimized by choice of permutation is then $s(D + U)$. Since $s(C + D)$ is independent of the permutation, the identity $D + U = \frac{1}{2}(C + D) + \frac{1}{2}(U - L)$ shows that an equivalent problem is to minimize $\frac{1}{2}s(U - L) = \sum_{k < m} \frac{1}{2}(c_{km} - c_{mk})$. Except for signs, the $\binom{n}{2}$ terms of this sum are the areas of all triangles determined in the first quadrant by the origin and any two points selected from the set with coordinates (a_i, b_i) , $i = 1, 2, \dots, n$. The choice of permutation therefore affects only the distribution of signs, and the minimum of the sum is attained on any permutation which attaches negative signs to all the non-zero terms. The permutation sought is therefore one for which b_{j_k}/a_{j_k} , $k = 1, 2, \dots, n$, is a non-increasing sequence, and is unique only if there is no pair of the points $\{(a_i, b_i)\}$ collinear with the origin.

III. *Comment by L. P. Prostanstus, Naval Electronics Laboratory Center, San Diego, California.* The solution to this problem has been published in J. G. Rau, *Minimizing a function of permutations of n integers*, Operations Research, 19 (1971), 237-240. (This paper describes some applications of this problem, such as minimizing expected cost of testing certain multi-component devices to find cause of failure. -Ed.)

Also solved by the proposer (see Comment), Peter de Buda, Sidney Heller, J. R. Kuttler, O. P. Lossers (Netherlands), L. P. Prostanstus (see Comment), F. G. Schmitt, Jr., and Nan-Shan Shou.

Derangements

E 2354 [1972, 394]. *Proposed by L. Carlitz and R. A. Scoville, Duke University*

Let $S = \{1, 2, \dots, n\}$ and let D_n denote the number of permutations of S with no fixed points (derangements). Let E_n denote the number of even permutations of S

with no fixed points. Show that

$$E_n = \binom{n}{2} D_{n-2} - (-1)^n (n-1), \quad n = 2, 3, \dots$$

Solution by Bob Prielipp and N. J. Kuenzi, University of Wisconsin, Oshkosh.
The following well-known formulas are given in the solution of Problem E 907:

$$(1) \quad D_n = n! \sum_{i=0}^n (-1)^i / i!, \quad (2) \quad E_n = \{D_n - (-1)^n (n-1)\} / 2.$$

(See pp. 687–688, December 1950 issue of this MONTHLY.)

Using (1) we have

$$\begin{aligned} \binom{n}{2} D_{n-2} &= n! \left\{ \sum_{i=0}^n (-1)^i / i! - (-1)^{n-1} / (n-1)! - (-1)^n / n! \right\} / 2 \\ &= \{D_n + (-1)^n (n-1)\} / 2. \end{aligned}$$

Using this last equation and (2) we have the desired result:

$$\binom{n}{2} D_{n-2} - (-1)^n (n-1) = \{D_n - (-1)^n (n-1)\} / 2 = E_n.$$

Also solved by Marc Berger, Problem Solving Group Berne (Switzerland), D. M. Bloom, Peter de Buda, J. Chone (France), M. G. Greening (Australia), Wells Johnson, Harry Lass, O. P. Lossers (Netherlands), L. E. Mattics, W. O. J. Moser, P. J. Murray, W. J. Sanchez, F. G. Schmitt, Jr., Allen Stenger, E. T. H. Wang, and the proposers.

An Early Jacobi

E 2357 [1972, 518]. *Proposed by M. D. Hirschhorn, Penicnik, Midlothian, Scotland*

Suppose that m and n are nonnegative integers and that x_0, x_1, \dots, x_m are distinct. Show that

$$\sum x_0^{k_0} \dots x_m^{k_m} = \sum_{i=0}^m \frac{x_i^{m+n}}{\prod_{j \neq i} (x_i - x_j)},$$

where the sum on the left-hand side is over all (k_0, k_1, \dots, k_m) with $k_i \geq 0$ and $k_0 + \dots + k_m = n$, and where the product is over all $j \neq i$.

Solution by M. G. Greening, University of New South Wales, Australia. Set $b_i = x_i^{m+n} / \prod_{j \neq i} (x_i - x_j)$. The left-hand side of the given identity is the coefficient of x^{-1} in the Laurent's expansion of $f(x) = x^{n+m} / \prod_{j=0}^m (x - x_j)$, where

$$|x| > \max \{x_0, x_1, \dots, x_m\},$$

as $f(x) = x^{n-1} \prod_{j=0}^m (1 - x_j/x)^{-1}$.

On the other hand, decomposition of $f(x)$ into partial fractions by the "cover-up" method gives:

$$f(x) = \sum_{i=0}^m b_i/(x - x_i) = \sum_{i=0}^m b_i x^{-1} (1 - x_i/x)^{-1}$$

and the coefficient of x^{-1} in this expansion is $\sum_{i=0}^m b_i$ which is the right-hand side of the identity.

Also solved by Peter Bundschuh (West Germany), Leonard Carlitz, J. E. Chance, R. J. Evans, S. H. Halton, A. C. Hindmarsh, M. S. Klamkin, O. P. Lossers (Netherlands), Helen M. Marston, M. R. Railkar (India), A. G. Shannon (Australia), Nan-Shan Shou, R. P. Soni, C. H. Yong, and the proposer.

Editor's comment. As pointed out by Klamkin and Railkar, this result is not new but is a result in the theory of alternants. For an alternate solution see *Theory of Determinants* by Muir and Metzler, paragraphs 333 and 335.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers—The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before July 31, 1973. Contributors (in the United States) who desire acknowledgement of receipt of their solutions are asked to enclose self-addressed, stamped postcards.

An asterisk () means neither the proposer nor the editors supplied a solution.*

5906. *Proposed by Gérard Letac, University of Clermont, France*

Let $x_0, x_1, \dots, x_t, \dots$ be independent random variables such that $P(x_t = n) = p_n < 1$ for all t and $n = 0, 1, 2, \dots$ with $\sum_{n=0}^{\infty} p_n = 1$; let q_n denote $P(x_t < n)$. Find the distribution of

$$q_{x_0} + p_{x_0} q_{x_1} + \dots + p_{x_0} p_{x_1} \dots p_{x_{t-1}} q_{x_t} + \dots$$

5907. *Proposed by J. C. Alexander, University of Maryland*

For $n \geq 1$, let S_n be the set of polynomials of the form

$$p(z) = z^n + a_{n-1}z^{n-1} + a_{n-2}z^{n-2} + \dots + a_1z + 1,$$

where a_1, a_2, \dots, a_{n-1} range through all complex numbers. What is the value of

$$M_n = \min_{p \in S_n} \left(\max_{|z|=1} |p(z)| \right)?$$

5908. *Proposed by M. L. Glasser, Battelle Memorial Institute*

Prove

$$\sum_{i,j,k,l} \Delta^{-4} \bigg/ \sum_{i,j,k,l} \Delta^{-3} = \pi^2/12,$$

where $\Delta = i^2 + j^2 + k^2 + l^2 + i + j + k + l + 1$ and the summations are over all non-negative integers.

5909. *Proposed by Gérard Letac, University of Clermont, France*

Let f be a continuous real function on some real, finite dimensional vector space E . For any base $b = (b_1, \dots, b_n)$ of E , denote by $E_b = \{z_1 b_1 + \dots + z_n b_n; z_i \in \mathbb{Z}, i = 1, \dots, n\}$, where \mathbb{Z} is the set of integers. Is it true that f is a bounded function when, for any base b , f restricted to E_b is bounded?

5910*. *Proposed by P. M. Eakin, University of Kentucky*

Suppose A and B are rings, n an integer and $X_1, \dots, X_n, Y_1, \dots, Y_n$ indeterminates. If $A[X_1, \dots, X_n] \cong B[Y_1, \dots, Y_n]$ and A is euclidean, then B is euclidean.

SOLUTIONS OF ADVANCED PROBLEMS

Powers of an Algebraic Real Number

5832 [1972, 93] *Proposed by Irwin Just, Bronx Community College*

Let n be an integer greater than one. Must there exist an algebraic real number r , of degree n such that for each positive integer m , $[r^m]$ is an odd integer? ($[x]$ is the greatest integer not exceeding x .)

Solution by Leonard Carlitz and Richard Scoville, Duke University. We shall prove the following more general result: *Let n, k be integers > 1 . Then there exists an algebraic number r of degree n such that*

$$[r^m] \equiv -1 \pmod{k} \quad (m = 1, 2, 3, \dots).$$

Proof. Let k be an integer > 1 and let $r_1, r_2, \dots, r_n, n \geq 2$, be the roots of

$$f(x) = \sum_{i=0}^n a_i x^i,$$

where the following conditions are satisfied:

- (1) $a_n = 1$.
- (2) all a_i are integers with $a_0 \equiv a_1 \equiv \dots \equiv a_{n-1} \equiv 0 \pmod{k}$.
- (3) f is irreducible.
- (4) all r_i are positive and $r_1 > r_2 > r_3 > \dots > r_n$.
- (5) $\sum_{i=2}^n r_i < 1$.

Then from (1) and (2) it follows that

$$L_j = \sum_{i=1}^n r_i^j$$

is an integer and that $L_j \equiv 0 \pmod{k}$ for every $j \geq 1$. Then, by (5),

$$L_j - 1 < r_1^j < L_j,$$

so we get $[r_1^j] \equiv -1 \pmod{k}$ ($j \geq 1$).

Now to solve the problem we must show the existence of a polynomial f_n satisfying (1)–(5). If (1), (2), (4) and (5) are satisfied, so is (3), since if a product is integral, the product must contain r_1 , by (5).

Let $f_2(x) = x^2 - Bkx + k$ where B is an integer chosen so large that (5) is satisfied. Hence (1)–(5) are satisfied.

Now assume f_{n-1} has been chosen to satisfy (1)–(5). We will see that an integer A can be chosen so large that

$$f_n(x) = x^n - f_{n-1}(A k x)$$

also satisfies (1) to (5): let r'_1 be the largest root of f_{n-1} . Then f_{n-1} changes sign $n-1$ times between 0 and $r'_1 + 1$. Hence $-f_{n-1}(A k x)$ changes sign $n-1$ times between 0 and $(r'_1 + 1)/A k$. If we choose A sufficiently large the effect on x^n is negligible and $f_n(x) = x^n - f_{n-1}(A k x)$ will change sign $n-1$ times between 0 and $(r'_1 + 1)/A k$. Furthermore $f_n((r'_1 + 1)/A k) < 0$, since $f_{n-1}(r'_1 + 1) > 0$. Thus f_n has $n-1$ roots between 0 and $(r'_1 + 1)/A k$ and another large positive root. Clearly by choosing A still larger, we can ensure (5).

We may also prove that if k is a fixed positive integer, there exists a real algebraic number r of degree 2 for which

$$[r^m] \equiv 1 \pmod{k}, \quad m = 1, 2, 3, \dots$$

Also solved by J. G. Mauldon, P. L. Montgomery, and the proposer.

Functions with a Limit

5833 [1972, 93]. Proposed by J. Bernard and G. Letac, University of Clermont, France

Let f be a continuous function on the positive real numbers such that $f(x) \leq f(nx)$ for any $x > 0$ and any integer $n > 0$. Prove that $\lim_{x \rightarrow \infty} f(x)$ exists ($\leq +\infty$).

Solution by William Wong, The City College, New York. Let

$$M = \lim_{r \rightarrow \infty} \sup_{x \geq r} f(x), \quad m = \lim_{r \rightarrow \infty} \inf_{x \geq r} f(x)$$

and suppose $M > m$. Choose k , $M > k > m$; then $f(a) > k$ for some a and by continuity, there exists $b > a$ such that for all $x \in [a, b]$, $f(x) > k$. Let $p = ab/(b-a)$. Now $x \geq p$ implies $x/a \geq x/b + 1$; thus $x/a \geq n \geq x/b$ for some positive integer n ; i.e., $a \leq x/n \leq b$. Hence $f(x/n) > k$, $f(x) > k$ for all $x \geq p$. This contradicts the definition of m and $m < M$.

Also solved by C. S. Allen, L. F. Bennett, P. R. Chernoff, R. A. Christiansen, L. E. Clarke (England), J. Cobb, R. O. Davies (England), L. Eifler, N. Felsing, P. Fisher (Netherlands), J. Fridy, L. Gerber, J. Gill, A. C. Hindmarsh, A. A. Jagers (Netherlands), D. R. King, J. Levy, O. P. Lossers (Netherlands), J. G. Mauldon, A. Meir, S. Monteferrante, P. L. Montgomery, M. Powderly, S. Rajnak, W. Snow, J. Sturm, R. K. Tamaki, Nguyen Xuan Uy, H. Van Evelghem (Belgium), F. I. Wright, K. L. Yocom, and the proposers.

Note. Van Evelghem and Mauldon relax the hypothesis by replacing n with u_n , $u_n \nearrow \infty$, $\limsup u_{n+1}/u_n = 1$.

Roots of Irreducible Polynomials of Prime Degree

5834 [1972, 93]. *Proposed by Irwin Just, Bronx Community College*

Let f be an irreducible seventh degree polynomial with rational coefficients and let S be a proper subset of the zeros of f . Can the sum of the elements of S be rational?

Solution by Stephen Pierce, University of Toronto. The answer is no, and we will prove this in the case that f has degree p , where p is any prime. Let G be the Galois group of f regarded as a subgroup of S_p . Since $p \mid |G|$, there is a p -cycle σ in G , say $\sigma = (123 \cdots p)$. Suppose $\lambda_1, \dots, \lambda_p$ are the roots of f and assume that

$$\lambda_{i_1} + \cdots + \lambda_{i_t} = r,$$

r a rational number. Apply $\sigma, \sigma^2, \dots, \sigma^{p-1}$ to this equation, thus obtaining p equations. Thus, there is a $p \times p$ matrix A with elements 0 and 1 such that $\lambda A = r$, where

$$\lambda = (\lambda_1, \dots, \lambda_p) \text{ and } r = (r, \dots, r).$$

Moreover, A is a circulant of degree p , with not all entries the same. Hence $\det(A) \neq 0$ and $\lambda = A^{-1}r$. Thus, $\lambda_1, \dots, \lambda_p$ are rational, a contradiction.

Also solved by L. Carlitz, A. A. Jagers (Netherlands), J. Tillman, and the proposer.

Convex Type Measurable Functions

5835 [1972, 94]. *Proposed by G. Letac, University of Clermont-Ferrand, Aubière, France*

Prove that the constants are the only measurable functions f on the positive real line such that for any positive x and y , $f(x+y)$ belongs to the interval spanned by $f(x)$ and $f(y)$.

Solution by J. G. Butler, University of Alberta. Suppose, on the contrary, that $f(x)$ is a measurable function on R^+ assuming at least two values a_1, a_2 , $a_1 < a_2$, at $x = x_1, x_2$, respectively. For $y > 0$, define

$$S_1(y) = \{x: 0 < x < y \text{ and } f(x) \leq a_1\}$$

$$S_2(y) = \{x: 0 < x < y \text{ and } f(x) \geq a_2\}.$$

If $x \in (0, \frac{1}{2}x_1) \sim S_1(\frac{1}{2}x_1)$, then $a_1 = f(x_1)$ lies between $f(x)$ and $f(x_1 - x)$ so that $x_1 - x \in S_1(x_1)$. Hence

$$S_1(x_1) \supset S_1(\frac{1}{2}x_1) \cup [x_1 - ((0, \frac{1}{2}x_1) \sim S_1(\frac{1}{2}x_1))]$$

and so $\text{meas}(S_1(x_1)) \geq \frac{1}{2}x_1$.

Now the hypotheses imply $f(qx) = f(x)$, with q a positive rational, so that for all $y > 0$ and all rationals $q < y/x_1$, we have $S_1(y) \supset q S_1(x_1)$. It follows that $\text{meas}(S_1(y)) \geq \frac{1}{2}y$. Similar reasoning yields $\text{meas}(S_2(y)) \geq \frac{1}{2}y$. Since $S_1(y)$, $S_2(y)$ are disjoint measurable sets, equality must hold.

It now follows, using additivity properties of measure that $\text{meas}(S_1(1) \cap M) = \frac{1}{2} \text{meas}(M)$ for each measurable subset M of $(0, 1)$. Taking $M = S_1(1)$ yields the desired contradiction.

Also solved by P. R. Chernoff, R. A. Christiansen, R. O. Davies (England), John Gill, Ralph Jones, Joel Levy, and the proposer.

Editor's note. Chernoff and the proposer note that the hypothesis of measurability is critical; for if g is a nonmeasurable additive function then $g(x)/x$ is not constant but $g(x+y)/(x+y)$ is a convex combination of $g(x)/x$ and $g(y)/y$, $x, y > 0$.

Minute Translates of a Measurable Function

5836 [1972, 94]. *Proposed by Eric Bedford and Michael Taylor, University of Michigan*

Let $f(x)$ be bounded and measurable on $(0, 1)$. Is it true that $\lim_{n \rightarrow \infty} f(x - 1/n) = f(x)$ almost everywhere? Prove, or provide a counterexample.

Solution by Amram Meir, University of Alberta. The statement is false. We shall prove that for a given $\varepsilon > 0$ there exists a bounded measurable $f(x)$, such that

$$(*) \quad \mu \left\{ x: \overline{\lim}_{n \rightarrow \infty} \left| f\left(x - \frac{1}{n}\right) - f(x) \right| = 1 \right\} > 1 - \varepsilon,$$

where $\mu\{\cdot\}$ denotes the Lebesgue measure. Let $x \in (0, 1)$, then x can be written in the form

$$x = \sum_{n=2}^{\infty} \frac{a_n}{n!}, \quad 0 \leq a_n \leq n-1,$$

where $a_n = [n!x] - n[(n-1)!x]$.

Let N be an integer such that $2^{-N} < \varepsilon$, let $n_k = 2^{N+k}$ ($k = 1, 2, \dots$). We define the subsets B_k of $(0, 1)$ by

$$B_k = \{x: a_{n_k} = 0\}, \quad k = 1, 2, \dots$$

Clearly, B_k is measurable and $\mu(B_k) = 2^{-N-k}$. Let now

$$A = \bigcap_{k=1}^{\infty} B_k^c,$$

where B_k^c denotes the complement of B_k in $(0, 1)$. A is measurable and $\mu A > 1 - \sum_{k=1}^{\infty} 2^{-N-k} > 1 - \varepsilon$. Now let $f(x) = 0$ for $x \in A$, $f(x) = 1$ for $x \in A^c$, and let x_0 be an arbitrary point in A . We have

$$x_0 = \sum_{n=2}^{\infty} \frac{a_n}{n!}, \quad a_{n_k} \neq 0 \quad \text{for } k = 1, 2, \dots.$$

It is easy to see that because of $a_{n_k} \leq n_k - 1$,

$$\frac{a_{n_k}}{(n_k)!} = \frac{1}{m_k}$$

with a suitable integer m_k , and that $m_k \rightarrow \infty$. Thus

$$x_0 - \frac{1}{m_k} \in B_k \subset A^c,$$

and so $f(x_0 - 1/m_k) = 1$, $k = 1, 2, \dots$, while $f(x_0) = 0$. We thus see that x_0 belongs to the set whose measure is estimated under (*). Since $\mu A > 1 - \varepsilon$, the inequality (*) follows.

Also solved by G. J. Butler, R. O. Davies (England), Douglas Lind, and the proposers.

Editor's note. Davies raises the question as to the consequence of replacing $1/n$ by some other null sequence. C. J. Neugebauer observes that the sequence $f_n(x) = f(x - 1/n)$ converges in mean to $f(x)$ and so some subsequence converges to $f(x)$ almost everywhere.

Inverses in Prime Rings

5837 [1972, 94]. *Proposed by I. N. Herstein and Susan Montgomery*

Cancelled. Duplicate: See solution 5797 [1972, 916].

Number of Non-isomorphic Groups of a Given Order

5838 [1972, 187]. *Proposed by R. B. Eggleton, University of Calgary*

Let $N(g)$ denote the number of isomorphism classes of abelian groups of order g . The equation $N(x) = n$ is solvable for $1 \leq n \leq 12$, and for infinitely many other natural numbers n , but there is no solution when $n = 13$. Show that there are infinitely many natural numbers n for which there is no solution.

Solution by Eric Rosenthal, student, Yale University. It is a consequence of the structure theorems for finite theorems for finite abelian groups that if q_1, q_2, \dots, q_k are distinct primes, then

$$N(q_1^{e_1} q_2^{e_2} \cdots q_k^{e_k}) = P(e_1) P(e_2) \cdots P(e_k),$$

where P denotes the partition function. So $N(x) = p$ has no solution if p is a prime not in the range of P .

It is known that the n th prime $p_n = O(n^2)$, [Problem E 2275 [1972, 89]] and that

$$P(n) \sim \frac{1}{4n\sqrt{3}} \exp(\pi\sqrt{2n/3}).$$

[Marshall Hall, *Combinatorial Theory*, Blaisdell, 1967, p. 40.]

So $p_n = o(P(n))$ and most primes are not partitions. If p is one of the infinitely many primes outside the range of P then $N(x) = p$ is insoluble.

Also solved by K. A. Beres, P. X. Gallagher, R. Harris & F. Oliva & R. Thrasher, C. V. Heuer & G. A. Heuer, P. L. Montgomery, and the proposer.

Generalization of the Power Function

5839 [1972, 187]. *Proposed by A. D. Ziebur, State University of New York at Binghamton*

The equation $\pi(x, y) = x^y$ defines a function from $R^+ \times R$ to R^+ (R is the set of real numbers, R^+ the set of positive reals) such that $\pi(x, n) = x^n$ when n is an integer, and $\pi(x, yz) = \pi(\pi(x, y), z)$. Is the power function the only function with these properties?

Solution by P. L. Montgomery, San Rafael, California. It is not the only function. Let $\{\ln 2, \ln 3, \ln 5, \dots\} \cup B$ be a Hamel basis for the real numbers. For each $x \in R^+$ let

$$\ln x = \sum_{p \text{ prime}} a_p \ln p + \sum_{b \in B} r_b b,$$

where the a 's and the r 's are rational numbers, almost all zero. Set

$$\sigma(x) = \sum_{p \text{ prime}} p^{a_p}.$$

Also set $\sigma(0) = 0$ and $\sigma(-x) = -\sigma(x)$ if $x > 0$. Then σ fixes all integers; also $\sigma(xy) = \sigma(x)\sigma(y)$ for all real x and y . It quickly follows that the function $\pi(x, y) = x^{\sigma(y)}$ satisfies the stated conditions, as well as $\pi(xy, z) = \pi(x, z)\pi(y, z)$, but is not the power function.

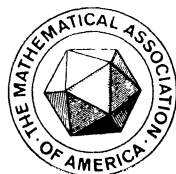
Also solved by O. P. Lossers (Netherlands) and the proposer.

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF
THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 80



NUMBER 5

CODEN: AMMYAE

CONTENTS

Notice	475
The Geometry of Connections R. S. MILLMAN AND ANN K. STEHNEY	475
An Introduction to Matroid Theory R. J. WILSON	500
Stochastic Equations and their Applications G. C. PAPANICOLAOU	526
The Main Crises S. BIRNBAUM	545

MATHEMATICAL NOTES

The Sign of the Bernoulli Numbers L. J. MORDELL	547
The Sign of the Bernoulli and Euler Numbers . . . L. CARLITZ AND R. SCOVILLE	548
Basically Bounded Sets and a Generalized Heine-Borel Theorem . NEIL HINDMAN	549

RESEARCH PROBLEMS

A Problem on Rational Functions A. K. PIZER	552
-------------------------------------------------------	-----

CLASSROOM NOTES

A Condition under which a Mapping is a Homeomorphism . . W. R. DERRICK	554
------------------------------------------------------------------------	-----

MATHEMATICAL EDUCATION

Independent Study for Undergraduates W. C. RAMALEY	555
--------------------------------------------------------------	-----

ELEMENTARY PROBLEMS AND SOLUTIONS	559
---------------------------------------------	-----

ADVANCED PROBLEMS AND SOLUTIONS	564
-------------------------------------------	-----

REVIEWS	568
-------------------	-----

NEWS AND NOTICES	585
----------------------------	-----

(Continued on inside cover)

MAY

1973

MATHEMATICAL ASSOCIATION OF AMERICA	586
November Meeting of the Ohio Section	586
November Meeting of the Seaway Section	587
The Fifty-sixth Annual Meeting of the Association	587
Announcement of the Walter B. Ford Lecture Fund	595
Academic Members Elected into the Association	597
Calendars of Future Meetings.	598

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss.

Backlog: Main Articles 12 months, Math. Notes 15 months, Research Problems 7 months, Classroom Notes 11 months, Math. Education 10 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to ALEX ROSENBERG, Department of Mathematics, Cornell University, Ithaca, N.Y. 14850; NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D.C. 20036.

HARLEY FLANDERS, *Editor*

ALEX ROSENBERG, *Editor-Elect*

ASSOCIATE EDITORS

JOSHUA BARLAZ
E. R. BERLEKAMP
JANE W. DI PAOLA
ROBERT GILMER
RICHARD GUY
RAOUL HAILPERN

J. G. HARVEY
ERIC S. LANGFORD
P. D. LAX
ARTHUR MATTUCK
M. W. POWNALL
GIAN-CARLO ROTA

SEYMOUR SCHUSTER
J. ARTHUR SEEBACH, JR.
E. P. STARKE
LYNN A. STEEN
JAMES WENDEL

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are \$12.50. For nonmembers the subscription price is \$18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June-July, August-September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices.

Copyright © The Mathematical Association of America (Incorporated), 1973

PRINTED IN THE UNITED STATES OF AMERICA

NOTICE

With the January 1974 issue the editorship of the MONTHLY will be taken over by Alex Rosenberg, Department of Mathematics, Cornell University, Ithaca, New York 14850, U.S.A. All manuscripts for main articles, and all editorial correspondence not concerned with previously submitted material, should be addressed to Ithaca from now on. The Statement of Policy (Vol. 76 (1969), p. 2) will of course remain in full force.

THE GEOMETRY OF CONNECTIONS

R. S. MILLMAN, Southern Illinois University and ANN K. STEHNEY, Wellesley College

Modern geometry was born when Riemann first separated the concept of geometry from the concept of space. His inaugural lecture at Göttingen, *On the Hypotheses which lie at the Foundation of Geometry* (1854), began: "As is well known, geometry presupposes the concept of space, as well as assuming the basic principles for constructions in space. . . . The relationship between these presuppositions is left in the dark." (From Spivak's translation [10], Vol. II. See also Smith [9].) The purpose of this paper is to see how 119 years of differential geometry and topology have shed light on this relationship. We shall first investigate the topological basis for modern differential geometry and then add some geometric structure to obtain the "principles for construction in space."

In Section 1 we define a differentiable manifold, which plays the role of "space" and provides points for our geometries. Riemann called this a continuous manifold or " n -fold extended quantity." Here we are in the realm of differential topology. To do geometry we need some structure which provides lines ("constructions in space"). The rest of the paper is concerned with three geometric structures, involving the use of calculus, which are motivated by the wish that lines (called geodesics) have properties analogous to those of straight lines in Euclidean space.

In Euclidean n -space \mathbb{R}^n , a straight line may be defined as either a curve α such that

$$\frac{d^2\alpha}{dt^2} = 0,$$

Richard S. Millman received his Cornell Ph. D. under H. C. Wang in 1971. He held an Assistant Professorship at Ithaca College before taking on his present position.

Ann K. Stehney received her Ph. D. in 1971 at SUNY, Stony Brook, under John A. Thorpe. She has been at Wellesley College since then. Both authors work in modern differential geometry.

Editor.

or a curve which represents the shortest path between points. The first condition means that the tangent vector to α is constant along α , or that the derivative of the tangent vector in the direction of the tangent vector (i.e., the acceleration) is zero.

Our first geometric structure (Section 2) is a linear connection or covariant differentiation which enables us to differentiate one vector field with respect to another and so mimic the first “definition” of a straight line in \mathbb{R}^n . We may then define a geodesic as a curve such that the covariant derivative of the tangent vector field in the direction of the tangent vector field is zero. Classically, a connection was defined by Christoffel (1869) to be a set of symbols $\{\Gamma_{ij}^k\}$ or Γ_{ij}^k . The modern viewpoint was finally formulated about 1950 by Koszul.

We shall see that the concept of a linear connection is equivalent to that of parallel translation of vectors along curves. In \mathbb{R}^n , this corresponds to moving the origin of a vector to any point along the curve, keeping its direction and magnitude fixed. The emphasis on parallel translation is due to T. Levi-Civita (1917). It was E. Cartan who saw the importance of parallel translation and who introduced a more generalized notion of “tangent space” in order to define a geometry. This made Cartan’s original work very difficult to read. However, after 1930 the notion of a fiber bundle was developed. By 1950 a theory of connections on fiber bundles emerged, due to Ehresmann, Chern, and others, and shed light on a great deal of Cartan’s work. We shall return to this in Section 4.

In our third section we define a Riemannian metric on a differentiable manifold. The choice of a Riemannian metric (there are many on any manifold) endows the manifold with much structure, for example that of a metric space in the usual sense. We may then ask that geodesics be curves which minimize distance. We shall see that every Riemannian metric also gives rise to a unique linear connection enjoying certain “nice” properties, and that curves which minimize distance are geodesics with respect to this connection (but not conversely). The notion of a Riemannian metric dates back to Riemann’s 1854 lecture. The association of a connection is due to Levi-Civita.

We finally return to the notion of parallel translation and show that it is equivalent to a unique path-lifting property on a principal fiber bundle. This supplies a geometric structure, also called a connection, on the fiber bundle. This approach is not only useful (see Chern [2]), but with today’s hindsight, also reasonable, because it includes Cartan’s theory of moving frames as a special case (see [10], Vol. II, Ch. 7). Spivak calls this final approach “not only more abstract, more elegant, and more incomprehensible, but also more general” ([10] Vol. II, pp. 8–16).

To summarize, we present the linear connection first as a basic, minimal geometric structure which gives “lines.” The existence of a Riemannian metric leads back naturally to a connection for which lines have nice length-minimizing properties similar to those of Euclidean space. A connection on a manifold is also equivalent to a connection on a particular principal fiber bundle.

The reader may notice that, among other important topics, we have omitted the theory of differential forms. We feel that the geometric content of forms is harder to grasp than the search for geodesics which we have undertaken. We refer the reader to Flanders [4] or Chern [2] for this topic.

Details of the ideas in this paper may be found in Hicks [5] and Spivak [10]. An excellent exposition of connections on fiber bundles has been given by Nomizu [8]. More detailed accounts may be found in Kobayashi and Nomizu [6], or Bishop and Crittenden [1].

1. Introduction to Differentiable Manifolds. A differentiable manifold locally enjoys the topological properties of Euclidean space. It provides a setting, more general than that of Euclidean space, in which the differentiability of functions is meaningful. The many definitions we shall give should be examined for their Euclidean analogue. Since the theorems in this section are versions of results from advanced calculus, we shall not attempt complete proofs.

Recall that a function f from (an open subset of) Euclidean n -space \mathbb{R}^n into m -space \mathbb{R}^m is differentiable (C^∞) if, writing $f(x) = (f_1(x), \dots, f_m(x))$, the coordinate functions f_i have continuous partial derivatives of all orders with respect to the coordinates of \mathbb{R}^n . If u_1, \dots, u_n denote the coordinate functions on \mathbb{R}^n and v_1, \dots, v_m denote those on \mathbb{R}^m , then $f_j = v_j \circ f$ and the Jacobian matrix at $x \in \mathbb{R}^n$ of f is the matrix

$$(1.1) \quad J_f(x) = \left[\frac{\partial f_j}{\partial u_i} \Big|_x \right] = \left[\frac{\partial (v_j \circ f)}{\partial u_i} \Big|_x \right]$$

of the partial derivatives at x of the coordinate functions of f . By the Inverse Function Theorem, f has a differentiable inverse in a neighborhood of x if and only if $J_f(x)$ is non-singular.

A **differentiable manifold** M of dimension n is a paracompact Hausdorff topological space and a collection $\{(U_\gamma, \phi_\gamma) \mid \gamma \in \Gamma\}$ satisfying

- (i) the U_γ are open sets in M with $M = \bigcup_{\gamma \in \Gamma} U_\gamma$ and each ϕ_γ is a homeomorphism from U_γ onto an open set in \mathbb{R}^n ,
- (ii) whenever $U_\gamma \cap U_\delta$ is non-empty, the homeomorphism $\phi_\delta \circ \phi_\gamma^{-1}$ from $\phi_\gamma(U_\gamma \cap U_\delta)$ onto $\phi_\delta(U_\gamma \cap U_\delta)$ is C^∞ as a map between Euclidean spaces, and
- (iii) the collection $\{(U_\gamma, \phi_\gamma)\}$ is maximal with respect to (i) and (ii), that is, it contains every (U, ϕ) satisfying (i) such that $\phi \circ \phi_\gamma^{-1}$ and $\phi_\gamma \circ \phi^{-1}$ are C^∞ for all $\gamma \in \Gamma$.

Since $\{U_\gamma\}$ is an open covering for M , each point of the manifold has a neighborhood homeomorphic to an open set in \mathbb{R}^n , a condition which is sometimes stated “ M is locally Euclidean.” The set U_γ is called a coordinate neighborhood for any of its points and the pair (U_γ, ϕ_γ) is called a chart for M . Condition (ii) distinguishes a differentiable manifold from a topological manifold; we shall refer to it as the compatibility of the charts. We shall use “ n -manifold” exclusively to mean a dif-

ferentiable manifold of dimension n . The collection $\{(U_\gamma, \phi_\gamma) \mid \gamma \in \Gamma\}$ will be called a **differentiable (or C^∞) structure** for M . Condition (iii) implies that there is a unique differentiable structure on M containing any collection of charts which satisfies (i) and (ii). In the examples which follow, we shall always mean the differentiable structure containing the given charts.

Examples. 1. \mathbb{R}^n itself is trivially an n -manifold, for example as determined by the single chart $(\mathbb{R}^n, \text{identity})$. Similarly any open set in \mathbb{R}^n is an n -manifold. A single chart will not suffice to define the manifold structure, however, unless M is homeomorphic to a (necessarily open) subset of \mathbb{R}^n .

2. S^1 , the unit sphere in \mathbb{R}^2 , is a differentiable 1-manifold. Writing $S^1 = \{(x, y) \mid x^2 + y^2 = 1\}$, we choose the U_γ 's to be open semicircles: U_1, \dots, U_4 are the intersections of S^1 with the right ($x > 0$), left ($x < 0$), upper ($y > 0$), and lower ($y < 0$) half-planes respectively (see Fig. 1a). We let ϕ_1 and ϕ_2 assign to a point its second coordinate and ϕ_3 and ϕ_4 assign its first. Each ϕ_γ is a homeomorphism from its domain U_γ onto the open interval $(-1, 1)$. We shall check the compatibility condition for $U_1 \cap U_4$ and leave the other cases to the reader (of course, $U_1 \cap U_2$ and $U_3 \cap U_4$ are empty). Let v be an element of $\phi_4(U_1 \cap U_4) = (0, 1)$. Then

$$\phi_1 \circ \phi_4^{-1}(v) = \phi_1(v, -[1 - v^2]^{\frac{1}{2}}) = -[1 - v^2]^{\frac{1}{2}}$$

(see Fig. 1b), and so $\phi_1 \circ \phi_4^{-1}$ is C^∞ homeomorphism from $(0, 1)$ onto $(-1, 0)$. Other choices for charts are possible—in fact, only two overlapping coordinate neighborhoods are needed.

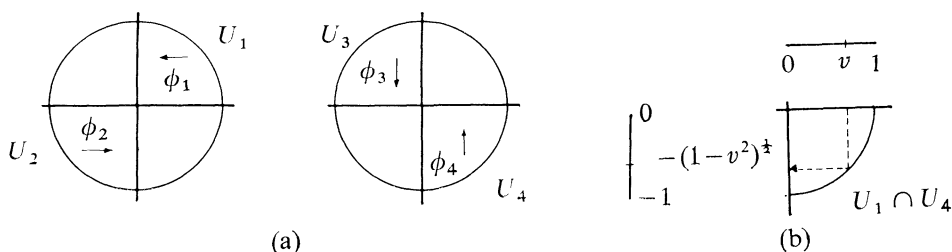


FIG. 1

3. S^2 , the unit sphere in \mathbb{R}^3 , is a differentiable 2-manifold. One choice of charts is analogous to the choice above for S^1 . The six coordinate neighborhoods are the hemispheres $x > 0$, $x < 0$, $y > 0$, $y < 0$, $z > 0$, and $z < 0$ (i.e., the intersections of S^2 with these open half-spaces) and the mappings ϕ_i are orthogonal projections onto the coordinate planes. For example, if U_1 is the hemisphere $x > 0$, then $\phi_1(x, y, z) = (y, z)$ and the image of ϕ_1 is the open unit disk in the (y, z) -plane. Thus in U_1 , a pair of numbers may be used as coordinates for points. The unique point in this hemisphere with coordinates (y, z) is $(+[1 - y^2 - z^2]^{\frac{1}{2}}, y, z)$. The reader may provide the details that the conditions for charts are satisfied.

4. The general linear group $G = GL(n, \mathbb{R})$ of all non-singular $n \times n$ matrices (under matrix product) is an n^2 -manifold. G is identified with an open set in \mathbb{R}^{n^2} via the map which strings out the matrix entries one row at a time:

$$[a_{ij}] \rightarrow (a_{11}, \dots, a_{1n}, a_{21}, \dots, a_{2n}, \dots, a_{nn}).$$

The determinate function on \mathbb{R}^{n^2} , a polynomial in its variables, is continuous, and G is the inverse image under "det" of the open set $\mathbb{R} - \{0\}$, and therefore open. Since $\det(G)$ is not connected, G is not.

Since differentiability is a local condition, we can extend that notion to continuous maps whose domain is a differentiable manifold and whose range is \mathbb{R} or any other differentiable manifold. The basic idea is to use the charts to move back to Euclidean space.

Let f be a real-valued function defined on an open subset V of M . Then f is called **differentiable** (or C^∞) at $m \in V$ if $f \circ \phi^{-1}$ is differentiable at $\phi(m)$ in \mathbb{R}^n for all charts (U, ϕ) for M at m . If V is open in M , f is called differentiable on V if it is differentiable at every $m \in V$. The set of all differentiable real-valued functions on V will be denoted $C^\infty(V)$.

Let M and N be manifolds, possibly of different dimensions, and f be a map from (an open set in) M into N . Let (U, ϕ) and (W, ψ) be charts for M and N at m and $f(m)$ respectively. Then f is called **differentiable** at m if $\psi \circ f \circ \phi^{-1}$ is differentiable at $\phi(m)$. This definition reduces to the one above for $N = \mathbb{R}$. It is not hard to show that the composition of differentiable maps is differentiable.

Examples. 1. If (U, ϕ) is a chart for the n -manifold M , then ϕ is C^∞ . If u_i is the i th coordinate function of \mathbb{R}^n , each $x_i = u_i \circ \phi$ is a real-valued function which is C^∞ on U . The function ϕ is in fact determined by the n real-valued functions x_1, \dots, x_n , called coordinate functions or a coordinate system for U .

2. If M_1 and M_2 are n_1 - and n_2 -dimensional manifolds respectively, the product space $M_1 \times M_2$ is a differentiable $(n_1 + n_2)$ -manifold in a natural way. The differentiable structure on $M_1 \times M_2$ is the one which makes the projection maps $\pi_i: M_1 \times M_2 \rightarrow M_i$ ($i = 1, 2$) onto the factor spaces differentiable.

3. Let $G = GL(n, \mathbb{R})$. The determinant function on G is C^∞ since it is the restriction of a polynomial to an open subset of \mathbb{R}^{n^2} . Furthermore, the group structure of G is consistent with its differentiable structure in the following sense. The functions $(A, B) \mapsto AB$ from $G \times G$ into G and $B \mapsto B^{-1}$ from G into G are C^∞ maps between manifolds. (Equivalently, the map $(A, B) \mapsto AB^{-1}$ from $G \times G$ into G is C^∞ .) Considering G as a subset of \mathbb{R}^{n^2} , the coordinate functions of \mathbb{R}^{n^2} provide local coordinates for G in a neighborhood of any point $A \in G$. The coordinates of AB (or B^{-1}) are sums of products of the coordinates of A and B (or rational functions of the coordinates of B alone), hence all partial derivatives of the maps exist. Similarly, for fixed $A \in G$, the maps L_A (left translation by A) and R_A (right translation by A), given by

$$L_A(B) = AB \text{ and } R_A(B) = BA,$$

are C^∞ maps from G onto itself.

$Gl(n, \mathbb{R})$ is an example of a **Lie group**, a group which is also a C^∞ manifold and whose operations $(g, h) \mapsto gh$ and $g \mapsto g^{-1}$ are C^∞ maps (from $G \times G$ into G and from G into G , respectively).

Just as a differentiable curve in Euclidean space has a tangent vector at each point, a differentiable manifold has at each point a collection of tangent vectors which form a linear space. If M is a submanifold of \mathbb{R}^q (a subset whose manifold structure is nicely related to that of \mathbb{R}), its tangent vectors at m may be pictured as the tangents at m to curves in \mathbb{R}^q whose images lie in M . (For a definition of "submanifold," see for example [5], p. 13.) Such a viewpoint is, however, not intrinsic, and we shall concentrate on the "directional derivative" aspect of a tangent vector. In \mathbb{R}^n , a tangent vector at a point assigns to each real-valued function a number, its derivative in that direction.

A **tangent vector** to M at m is a function X_m from $C^\infty(M)$ to \mathbb{R} which satisfies

- (i) (linearity) $X_m(af + bg) = aX_m(f) + bX_m(g)$, and
- (ii) (product rule) $X_m(fg) = f(m)X_m(g) + g(m)X_m(f)$,

for all real numbers a and b and all f and g in $C^\infty(M)$. For $M = \mathbb{R}^n$, it may be shown using Taylor's theorem (see [5]) that any X_m satisfying (i) and (ii) may be expressed uniquely as

$$\sum_i a_i \frac{\partial}{\partial u_i} \Big|_m$$

where the a_i 's are constants and the summation ranges over $i = 1, \dots, n$. In fact, $a_i = X_m(u_i)$. Thus the correspondence $X_m \leftrightarrow (a_1, \dots, a_n)$ gives an isomorphism of the set of tangent vectors to \mathbb{R}^n at the point m and the vector space \mathbb{R}^n itself. Since m is also an n -tuple, our X_m is a tangent vector in the classical sense, namely, (m, a_1, \dots, a_n) .

It is easily verified that for any constant function c , $X_m(c) = 0$. In addition, $X_m(f)$ depends on f only in a neighborhood of m , so that if functions f and g agree on an open set containing m , then $X_m(f) = X_m(g)$. Therefore $X_m(f)$ makes sense when f is defined only on an open set containing m (see [10], Vol. 1, p. 4-2).

The **tangent space** to M at m , denoted $T_m M$, is the set of tangent vectors to M at m , with the usual addition and scalar multiplication for real-valued functions:

$$(X + Y)_m(f) = X_m(f) + Y_m(f)$$

$$(aX)_m(f) = aX_m(f)$$

for $X_m, Y_m \in T_m M$, $f \in C^\infty(M)$, and $a \in \mathbb{R}$. The tangent space has been defined independently of any coordinate system for M near m ; however, the choice of a coordinate system provides a basis for $T_m M$. Let (U, ϕ) be a chart for M with $m \in U$,

and let $x_i = u_i \circ \phi$ for $i = 1, \dots, n$. Define the function $\partial/\partial x_i|_m$ from $C^\infty(M)$ to \mathbb{R} by

$$\frac{\partial}{\partial x_i}\bigg|_m(f) = \frac{\partial}{\partial u_i}\bigg|_{\phi(m)}(f \circ \phi^{-1})$$

for $f \in C^\infty(M)$. The right-hand side is the partial derivative of a real-valued function on \mathbb{R}^n . The linearity and product rule for partial differentiation imply that $\partial/\partial x_i|_m$ is a tangent vector to M at m . We may write

$$\frac{\partial f}{\partial x_i}\bigg|_m \text{ for } \frac{\partial}{\partial x_i}\bigg|_m(f).$$

The notation $\partial/\partial x_i$ suggests derivative in the x_i -direction. This interpretation is consistent with the identification of U and $\phi(U)$ in \mathbb{R}^n .

LEMMA. *The set of tangent vectors $\{\partial/\partial x_i|_m\}$ is linearly independent in $T_m M$.*

Proof. Assume that some linear combination $\sum_i a_i(\partial/\partial x_i)|_m$ is zero in $T_m M$. (All sums are taken over the range $i = 1, \dots, n$.) Then for every C^∞ function f , $\sum_i a_i(\partial f/\partial x_i)|_m = 0$. In particular, for $x_j = u_j \circ \phi$ ($1 \leq j \leq n$),

$$0 = \sum_i a_i \frac{\partial x_j}{\partial x_i}\bigg|_m = \sum_i a_i \frac{\partial u_j}{\partial u_i}\bigg|_{\phi(m)} = a_j$$

since $\partial u_j/\partial u_i|_{\phi(m)} = \delta_{ij}$ (Kronecker delta). \parallel

The proof that $\{\partial/\partial x_i|_m\}$ spans $T_m M$ is too long to be included here (see [5], p. 7). Once that is established, we see that $\{\partial/\partial x_i|_m\}$ is a basis for $T_m M$ and the tangent space at any point has the same dimension as the manifold. With respect to this basis, the tangent vector X_m at m is expressed

$$(1.2) \quad X_m = \sum_i X_m(x_i) \frac{\partial}{\partial x_i}\bigg|_m.$$

This method actually provides a basis for $T_q M$ for every $q \in U$. If $\{a_i\}$ is a collection of real-valued functions on U , $\sum_i a_i(\partial/\partial x_i)$ assigns to each $q \in U$ a tangent vector,

$$\sum_i a_i(q) \frac{\partial}{\partial x_i}\bigg|_q.$$

A **vector field** X on M is a function which assigns to each $m \in M$ an element X_m in $T_m M$. Every vector field takes the set of real-valued functions on M into itself: for $f: M \rightarrow \mathbb{R}$, Xf is the function from M to \mathbb{R} given by $(Xf)(m) = X_m(f)$. We are usually interested only in those which take $C^\infty(M)$ into itself. This property may be checked locally, for if X is a vector field on M , the following are equivalent:

- (i) for every $f \in C^\infty(M)$, Xf is in $C^\infty(M)$, and
- (ii) for every chart (U, ϕ) on M , the functions $a_i: U \rightarrow \mathbb{R}$ defined by

$$X_m = \sum_i a_i(m) \left. \frac{\partial}{\partial x_i} \right|_m \text{ are } C^\infty.$$

That is, in U , X is a linear combination of the vector fields $\partial/\partial x_i$ with coefficients in $C^\infty(U)$.

If X has these properties, it is called a **differentiable** (or C^∞) **vector field**. $\mathfrak{X}(M)$ will denote the vector space of all C^∞ vector fields on M . Note that if f is in $C^\infty(M)$ and X is in $\mathfrak{X}(M)$, then fX is in $\mathfrak{X}(M)$, whereas Xf is in $C^\infty(M)$.

If $f: M \rightarrow N$ is a C^∞ map between manifolds, any function g in $C^\infty(M)$ may be composed with f to yield a function $g \circ f$ in $C^\infty(M)$. This provides a mapping f_* , the **differential** of f , from the set of tangent vectors to M into the set of tangent vectors to N . It is perhaps easiest to think of f_* pulling functions on N back to M for differentiation. If X_m is in $T_m M$, we specify $f_*(X_m) \in T_{f(m)} N$ by giving its value on any $g \in C^\infty(N)$:

$$[f_*(X_m)](g) = X_m(g \circ f).$$

It is easy to check that for each $m \in M$, f_* is a linear map from $T_m M$ into $T_{f(m)} N$. When necessary to avoid confusion, we will write $f_*|_m$ for the restriction of f_* to $T_m M$.

If charts (U, ϕ) and (V, ψ) are chosen for M near m and for N near $f(m)$ respectively, the differential of f at m may be expressed as a matrix with respect to the determined bases for $T_m M$ and $T_{f(m)} N$. Let $\{x_1, \dots, x_{n_1}\}$ and $\{y_1, \dots, y_{n_2}\}$ be the local coordinates on M and N provided by ϕ and ψ . Then

$$\begin{aligned} f_* \left(\left. \frac{\partial}{\partial x_i} \right|_m \right) &= \sum_j \left[f_* \left. \frac{\partial}{\partial x_i} \right|_m \right] (y_j) \left. \frac{\partial}{\partial y_j} \right|_{f(m)} \\ &= \sum_j \left. \frac{\partial(y_j \circ f)}{\partial x_i} \right|_m \left. \frac{\partial}{\partial y_j} \right|_{f(m)}. \end{aligned}$$

With respect to the bases $\{\partial/\partial x_i|_m\}$ and $\{\partial/\partial y_j|_{f(m)}\}$, f_* is represented by the transpose of the matrix

$$J_f(m) = \left[\left. \frac{\partial(y_i \circ f)}{\partial x_j} \right|_m \right].$$

$J_f(m)$ is called the **Jacobian** of f at m (compare with Equation 1.1).

The differentials of compositions of C^∞ maps obey a chain rule. If $f: M \rightarrow N$ and $g: N \rightarrow L$ are C^∞ , then for $m \in M$, $(g \circ f)_*|_m = g_*|_{f(m)} \circ f_*|_m$. Similarly, $J_{g \circ f}(m) = J_g(f(m)) \circ J_f(m)$.

By a **curve** in M , we shall mean a C^∞ map $\alpha: I \rightarrow M$, where $I \subset \mathbb{R}$ is an interval. Notice that distinct curves α_1 and α_2 may have the same image in M , for example $\alpha_1(t) = (t, t)$ and $\alpha_2(t) = (t^3, t^3)$, where $M = \mathbb{R}^2$. The **tangent** to α at the point $t = t_0$, denoted $\dot{\alpha}(t_0)$, is the vector $\alpha_*(d/dt|_{t_0})$ in $T_{\alpha(t_0)} M$, where d/dt is the usual differentiation operator for functions of a real variable. We shall often consider

curves with a compact domain. The tangent to α at an endpoint is then understood to be the tangent to (any) C^∞ extension of α to a larger interval.

An **integral curve** of a vector field X is a curve α for which $\dot{\alpha}(t) = X_{\alpha(t)}$ for all t . For every $m \in M$, a C^∞ vector field X on M has an integral curve α defined on some interval $(-\varepsilon, \varepsilon)$ with $\alpha(0) = m$. It is unique in the sense that if $\beta: (-\delta, \delta) \rightarrow M$ is also an integral curve of X with $\beta(0) = m$, then $\alpha(t) = \beta(t)$ for all $t \in (-\varepsilon, \varepsilon) \cap (-\delta, \delta)$. Expressed in local coordinates for M near m , the condition $\dot{\alpha}(t) = X_{\alpha(t)}$ is a system of ordinary differential equations, and the results follow from the fundamental existence and uniqueness theorem for solutions to such systems.

In the same way that homeomorphic spaces are equivalent topologically, two differentiable manifolds are considered equivalent if they are **diffeomorphic**, that is, homeomorphic via a C^∞ map whose inverse is also C^∞ . They necessarily have the same dimension.

Differentiable manifolds are the “spaces,” the point-sets, for differential geometry.

2. Geometry as a linear connection. In this section we shall define a geometric structure on M , motivated by the fact that in \mathbb{R}^n we may differentiate one vector field with respect to another. In this formulation a “straight line” has the property that the derivative of its tangent vector field with respect to itself is zero. The references for this section are [5], p. 56, and [10], Vol. II, Ch. 6.

A **linear connection** (or **covariant derivative**) is an assignment of a vector field $\nabla_X Y$ to each pair $X, Y \in \mathfrak{X}(M)$ such that for all $X, Y, Z \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$,

- (i) $\nabla_{X+Y} Z = \nabla_X Z + \nabla_Y Z$,
- (ii) $\nabla_{fX} Y = f \nabla_X Y$,
- (iii) $\nabla_X (Y + Z) = \nabla_X Y + \nabla_X Z$, and
- (iv) $\nabla_X (fY) = f \nabla_X Y + (Xf)Y$.

Note that by conditions (i) and (ii), $\nabla_X Y$ at m depends only on X at m and not on X in a neighborhood of m . Therefore $\nabla_{X_m} Y$ makes sense and is an element of $T_m M$. The fourth condition makes sense because $fY \in \mathfrak{X}(M)$, whereas $(Xf)Y$ is just the product of the C^∞ vector field Y and the C^∞ function Xf . It is important to note that there are many different linear connections on a given manifold, i.e., a space may have different geometries on it. We shall see an example of this in the next section.

A curve $\alpha: I \rightarrow M$ is a **geodesic** if $\nabla_{\dot{\alpha}(t)} \dot{\alpha}(t) = 0$. At first glance this definition makes no sense because $\dot{\alpha}$ assigns a tangent vector only to the points $\alpha(I) \subset M$, not to every point in M . With a little work it can be shown that if X is any vector field with $X_{\alpha(t)} = \dot{\alpha}(t)$, then $\nabla_X X(\alpha(t))$ is independent of the extension X . By $\nabla_{\dot{\alpha}(t)} \dot{\alpha}(t)$, we therefore understand $\nabla_X X(\alpha(t))$ for any such X .

There is a technical problem here in that if $\dot{\alpha}$ is zero at some point, it may not be possible to find an extension X of $\dot{\alpha}$. On the other hand, we are only interested in geodesics and it can be shown that, for a geodesic α , if $\dot{\alpha}(t) = 0$ for some t , then α is a constant curve, and so we shall ignore this technical point.

Example. The standard connection on \mathbb{R}^n . If $Y \in \mathfrak{X}(\mathbb{R}^n)$, we may write $Y_p = \sum_j f_j(p) \partial/\partial u_j|_p$ for some $f_j \in C^\infty(\mathbb{R}^n)$, since $\{\partial/\partial u_j|_p\}$ is a basis for $T_p\mathbb{R}^n$ at all $p \in \mathbb{R}^n$. We define $\nabla_X Y(p) = \sum_j [X f_j](p) \partial/\partial u_j|_p$. In particular,

$$\nabla_{\partial/\partial u_i} \partial/\partial u_j = 0.$$

If $\alpha(t) = (\alpha_1(t), \dots, \alpha_n(t))$, then it can be shown that

$$\nabla_{\dot{\alpha}} \dot{\alpha} = \sum_i \frac{d^2 \alpha_j}{dt^2} \frac{\partial}{\partial u_i}$$

and so $\nabla_{\dot{\alpha}} \dot{\alpha}$ is a measure of the acceleration of α . It is also immediate that the geodesics are precisely the usual straight lines on \mathbb{R}^n . For instance, if $\alpha: \mathbb{R} \rightarrow \mathbb{R}^2$ is the curve $\alpha(t) = (\cos t, \sin t)$, then

$$\dot{\alpha}(t) = (-\sin t) \frac{\partial}{\partial u_1} + (\cos t) \frac{\partial}{\partial u_2}.$$

If $X_{(u_1, u_2)} = -u_2(\partial/\partial u_1) + u_1(\partial/\partial u_2)$, then $X_{\alpha(t)} = X_{(\cos t, \sin t)} = \dot{\alpha}(t)$. Hence

$$\begin{aligned} \nabla_{\dot{\alpha}(t)} \dot{\alpha}(t) &= \nabla_X X(\alpha(t)) \\ &= \left[-u_2 \frac{\partial}{\partial u_1} + u_1 \frac{\partial}{\partial u_2} \right] (-u_2) \frac{\partial}{\partial u_1} + \left[-u_2 \frac{\partial}{\partial u_1} + u_1 \frac{\partial}{\partial u_2} \right] (u_1) \frac{\partial}{\partial u_2} \\ &= -u_1 \frac{\partial}{\partial u_1} - u_2 \frac{\partial}{\partial u_2} \end{aligned}$$

which is non-zero, as we know, because a circle is not a geodesic in \mathbb{R}^2 (Fig. 2).

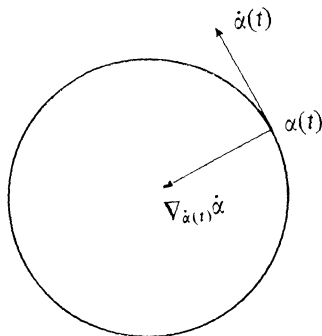


FIG. 2.

Geodesics starting at a given point $m \in M$ always exist, although they may not be defined on a very large interval. In fact, the following theorem says that an initial tangent may also be specified.

THEOREM 2.1. *Let m be a point of M . For each $X_m \in T_m M$, there exists a geodesic $\alpha: (-\varepsilon, \varepsilon) \rightarrow M$ (for some $\varepsilon > 0$) such that $\alpha(0) = m$ and $\dot{\alpha}(0) = X_m$.*

This "existence and uniqueness" theorem is so similar, in spirit and in proof, to Theorem 2.2 that we shall postpone a discussion of its proof.

A **vector field along a curve** $\alpha: I \rightarrow M$ is a function Y which assigns to each $t \in I$ a vector $Y_{\alpha(t)}$ in $T_{\alpha(t)}M$ such that for all $f \in C^\infty(M)$, $t \rightarrow Y_{\alpha(t)}f$ is a C^∞ real-valued function on I . If Y is a vector field along α , then Y is **parallel along** α provided $\nabla_{\dot{\alpha}(t)} Y = 0$ for all $t \in I$. For the standard connection on \mathbb{R}^n , $Y = \sum_i f_i (\partial/\partial u_i)$ is parallel along α if and only if $f_i \circ \alpha$ is constant for each i . For example, the basis vector fields $\{\partial/\partial u_i\}$ are parallel along any curve.

The next two theorems will be essential in the last section. The first is an existence theorem.

THEOREM 2.2. *Let α be a curve in M and $m = \alpha(0)$. For each X_m in $T_m M$, there is a unique vector field Y defined along α such that Y is parallel along α and $Y_m = X_m$.*

Proof (sketch). We do this locally. In a coordinate neighborhood U of $\alpha(0)$, we may write $\alpha(t) = (\alpha_1(t), \dots, \alpha_n(t))$, i.e., $\alpha_i = x_i \circ \alpha$, and $\dot{\alpha}(t) = \sum_i (d\alpha_i/dt)(\partial/\partial x_i)$ by Equation 1.2. We may also write

$$\nabla_{\partial/\partial x_i} \frac{\partial}{\partial x_j} = \sum_k \Gamma_{ij}^k \frac{\partial}{\partial x_k}$$

for some n^3 functions $\Gamma_{ij}^k \in C^\infty(U)$. If $Y = \sum_j f_j (\partial/\partial x_j)$, then $\nabla_{\dot{\alpha}(t)} Y = 0$ if and only if

$$\sum_{i,j} \left[\frac{d\alpha_i}{dt} \frac{\partial f_j}{\partial x_i} \frac{\partial}{\partial x_j} + \sum_k f_j \frac{d\alpha_i}{dt} \Gamma_{ij}^k \frac{\partial}{\partial x_k} \right] = 0$$

along α , or (evaluating the left-hand side on x_k) for each k

$$\sum_i \left[\frac{d\alpha_i}{dt} \frac{\partial f_k}{\partial x_i} + \sum_j f_j \frac{d\alpha_i}{dt} \Gamma_{ij}^k \right] = 0$$

or

$$(2.1) \quad \frac{d(f_k \circ \alpha)}{dt} + \sum_{i,j} (f_j \circ \alpha) \frac{d\alpha_i}{dt} \Gamma_{ij}^k \circ \alpha = 0$$

subject to the condition that the $f_k \circ \alpha(0)$ are given to be the components of X_m . We now invoke the fundamental existence and uniqueness theorem for ordinary differential equations where Γ_{ij}^k and $d\alpha_i/dt$ are known and $f_k \circ \alpha$ are unknown. This will give us $(f_k \circ \alpha)(t)$ for small t . Continuing in this way and using the compactness of I , we obtain a well-defined vector field Y along α . \parallel

The Γ_{ij}^k introduced in this proof are the classical Christoffel symbols. In terms of them, the condition that α be a geodesic is

$$\frac{d^2 \alpha_k}{dt^2} + \sum_{i,j} \Gamma_{ij}^k \frac{d\alpha_i}{dt} \frac{d\alpha_j}{dt} = 0.$$

This system has a unique solution (for small t) with given $\{\alpha_i(0)\}$ and $\{d\alpha_i/dt|_0\}$, that is, with given $\alpha(0)$ and $\dot{\alpha}(0)$, providing a proof of Theorem 2.1.

Returning to the parallel vector fields along α and the notation of Theorem 2.2, let $\tau_{\alpha(t)}: T_{\alpha(0)}M \rightarrow T_{\alpha(t)}M$ be defined by $\tau_{\alpha(t)}(X_m) = Y_{\alpha(t)}$. The map τ_α is called **parallel translation** along α , and an examination of the proof of Theorem 2.2. yields:

THEOREM 2.3. $\tau_{\alpha(t)}$ is a (vector space) isomorphism of $T_{\alpha(0)}M$ onto $T_{\alpha(t)}M$.

In \mathbb{R}^n with the standard connection, τ_α is independent of α , essentially because $\Gamma_{ij}^k = 0$ in Equation 2.1. In fact,

$$\tau_{\alpha(t)} \left(\sum_i a_i \frac{\partial}{\partial x_i} \Big|_{\alpha(0)} \right) = \sum_i a_i \frac{\partial}{\partial x_i} \Big|_{\alpha(t)}.$$

This is not the case in most other geometries and this is what makes \mathbb{R}^n so special. If in \mathbb{R}^n a vector is parallel translated around a closed curve, it comes back to itself (i.e., let α be a curve in \mathbb{R}^n such that $\alpha(0) = \alpha(1) = m$; then $\tau_{\alpha(1)}: T_m\mathbb{R}^n \rightarrow T_m\mathbb{R}^n$ is the identity). This phenomenon is because \mathbb{R}^n has zero "curvature." In this paper we are only concerned with connections and other structures which give connections and so we shall omit curvature although it is an important concept in differential geometry. Curvature is the theme of the second volume of Spivak [10] and we refer the reader there.

It may be surprising that parallel translation is just a global version of covariant differentiation. We mean this in the following sense.

THEOREM 2.4. Giving a linear connection on a manifold M is equivalent to giving for each curve α and each t , a vector space isomorphism $\tau_{\alpha(t)}: T_{\alpha(0)}M \rightarrow T_{\alpha(t)}M$ such that for every $X_{\alpha(0)} \in T_{\alpha(0)}M$

- (i) the assignment $t \rightarrow \tau_{\alpha(t)}X_{\alpha(0)}$ is C^∞ and
- (ii) for all $A \in GL(n, \mathbb{R})$,

$$\tau_{\alpha(t)}(AX_{\alpha(0)}) = A(\tau_{\alpha(t)}X_{\alpha(0)}).$$

Note. If x_1, \dots, x_n form a coordinate system in a neighborhood of $\alpha(t_0)$, we may write

$$\tau_{\alpha(t)}X_{\alpha(0)} = \sum_i a_i(t) \frac{\partial}{\partial x_i} \Big|_{\alpha(t)}$$

for t in a neighborhood of t_0 , where the a_i 's are real-valued functions. Condition (i) means that the a_i 's must be C^∞ on this neighborhood.

Proof. We saw in Theorem 2.3 that every linear connection gives rise to parallel translation. If we have parallel translation, then given X_m and $Y \in \mathfrak{X}(M)$, we define $(\nabla_X Y)(m)$ as follows. Let α be any integral curve of X_m with $\alpha(0) = m$ and set

$$(2.2) \quad (\nabla_X Y)(m) = \lim_{t \rightarrow 0} \frac{1}{t} (\tau_{\alpha(t)}^{-1} Y_{\alpha(t)} - Y_m),$$

where the limit is in $T_m M$ which is isomorphic to \mathbb{R}^n . We refer to [10], Vol. II, pp. 6–11 for details. \parallel

Using parallel translation we may compare the tangent spaces at any two points of M which may be joined by a curve α . Explicitly, we may define $\prod_{\alpha(t_0)}^{\alpha(t_1)} = \tau_{\alpha(t_1)} \circ \tau_{\alpha(t_0)}^{-1}: T_{\alpha(t_0)} M \rightarrow T_{\alpha(t_1)} M$ (which is of course an isomorphism) which allows us to compare the tangent spaces. In general, this isomorphism will depend on α .

Formula 2.2 means that covariant differentiation really measures how much the given vector field deviates from being parallel.

Because of Theorem 2.4, we can say that parallel translation is a geometric structure. We shall use this interpretation of geometric structure to motivate the definition of a connection on a fiber bundle in the last section.

3. Geometry as a Riemannian metric. In this section we shall add to a manifold M a structure (Riemannian metric) which makes M into a metric space. This structure will also lead us to a natural linear connection on M and we shall then see whether the set of geodesics coincides with the set of length-minimizing curves, as it does on \mathbb{R}^n .

This theory is appealing because the notion of distance is a familiar one. It will necessarily be more complicated than the same notion for Euclidean space, where things are measured by comparison with straight lines, and in fact the assignment of a length to a single line suffices. In his 1854 lecture, Riemann pointed out that there is no reason why the length of a line should be assumed to be independent of its position. He proposed measuring infinitesimals (for our purposes, tangent vectors, representing velocity) and integrating over a curve to find its length. The metric on the manifold is finally given in terms of lengths of curves. The increased generality of this method should be obvious; its simplicity lies in our use of calculus to investigate geometry.

A **Riemannian metric** (or inner product) g for M is the assignment of a positive-definite inner product g_m to each tangent space $T_m M$, which is differentiable in the sense that for any C^∞ vector fields X and Y on M , the function $m \mapsto g_m(X_m, Y_m)$ from M to \mathbb{R} is C^∞ . A **Riemannian manifold** is a differentiable manifold M provided with a Riemannian metric. Every differentiable manifold may be endowed with a Riemannian metric, for example, by using a C^∞ partition of unity to piece together arbitrary C^∞ metrics on the coordinate neighborhoods ([5], p. 85).

The length of a vector X_m in the inner product space $T_m M$ is given as usual by $\|X_m\| = [g_m(X_m, X_m)]^{\frac{1}{2}}$.

Let α be a smooth curve in M with domain $[a, b]$. The **length** of α is defined to be

$$L_a^b(\alpha) = \int_a^b \|\dot{\alpha}(t)\| dt.$$

The integral exists since $\|\dot{\alpha}(t)\|$ is a continuous function of t . If α is a piecewise C^∞ curve (α is continuous and its domain is the finite union of intervals in which α is C^∞) its length is defined to be the sum of the lengths of the C^∞ pieces. If there is no confusion, we shall write $L(\alpha)$ for $L_a^b(\alpha)$.

If M is a connected (therefore path-connected) Riemannian manifold, there is at least one piecewise C^∞ curve (in fact, at least one C^∞ curve) joining any two points of M . We define a distance function $d: M \times M \rightarrow \mathbb{R}$ by

$$d(m, q) = \inf\{L(\alpha) \mid \alpha \text{ is a piecewise } C^\infty \text{ curve in } M \text{ with endpoints } m \text{ and } q\}$$

Clearly $d(m, q) \geq 0$ and $d(m, m) = 0$. It is also true that $d(m, q) > 0$ unless $m = q$, and that for any third point $r \in M$, $d(m, q) \leq d(m, r) + d(r, q)$ (the triangle inequality). The function d is thus a metric (in the usual sense) for M , determined by g . Different g 's will yield different metrics for M , but such metric topologies for M are always the original manifold topology ([5], p. 70).

Examples: 1. As usual, let $\{u_i\}$ be the coordinate functions of Euclidean space. Define (the standard) Riemannian metric on \mathbb{R}^n by

$$g_m \left(\left. \frac{\partial}{\partial u_i} \right|_m, \left. \frac{\partial}{\partial u_j} \right|_m \right) = \delta_{ij}.$$

In particular, the vectors $\partial/\partial u_1|_m$ and $\partial/\partial u_2|_m$ form an orthonormal basis for $T_m\mathbb{R}^2$. Let $\alpha = (\alpha_1, \alpha_2)$ be a curve in \mathbb{R}^2 . By equation 1.2 and the definition of $\dot{\alpha}$,

$$\dot{\alpha} = \frac{d}{dt}(u_1 \circ \alpha) \frac{\partial}{\partial u_1} + \frac{d}{dt}(u_2 \circ \alpha) \frac{\partial}{\partial u_2}$$

and

$$\|\dot{\alpha}\| = \left[\left(\frac{d\alpha_1}{dt} \right)^2 + \left(\frac{d\alpha_2}{dt} \right)^2 \right]^{\frac{1}{2}}.$$

Let us compute the lengths of the straight and semicircular segments which join $(-1, 0)$ to $(1, 0)$. In the first case, $\alpha(t) = (2t - 1, 0)$ for $t \in [0, 1]$. Since $\dot{\alpha} = 2(\partial/\partial u_1)$, we have $\|\dot{\alpha}(t)\| = 2$ for all t , and $L(\alpha) = 2$. In the second case, $\alpha(t) = (-\cos \pi t, \sin \pi t)$ for $t \in [0, 1]$. Then

$$\dot{\alpha} = \pi \sin \pi t \frac{\partial}{\partial u_1} + \pi \cos \pi t \frac{\partial}{\partial u_2},$$

so that $\|\dot{\alpha}(t)\| = \pi$ for all t and $L(\alpha) = \pi$. This inner product gives the usual metric structure for \mathbb{R}^2 and may be restricted to any open subset of \mathbb{R}^2 to give its usual metric structure.

2. Let H be the open upper half-plane in \mathbb{R}^2 with its usual subspace manifold structure. Define g by

$$g_{(u_1, u_2)} \left(\frac{\partial}{\partial u_i}, \frac{\partial}{\partial u_j} \right) = \frac{\delta_{ij}}{u_2^2}.$$

For $\varepsilon > 0$, let us compute the length of the curve $\alpha_\varepsilon: [\varepsilon, 1] \rightarrow H$ given by $\alpha_\varepsilon(t) = (0, t)$. Since $\dot{\alpha}_\varepsilon = \partial/\partial u_2$, we have $\|\dot{\alpha}_\varepsilon(t)\| = 1/t$. Therefore

$$L(\alpha_\varepsilon) = \ln t \Big|_\varepsilon^1 = -\ln \varepsilon.$$

As ε approaches zero, $L(\alpha_\varepsilon)$ obviously becomes infinite.

3. Let $h: S^2 \rightarrow \mathbb{R}^3$ be the inclusion map and let g^3 denote the standard metric on \mathbb{R}^3 . We define the **induced metric** g on S^2 by

$$g(X, Y) = g^3(h_*X, h_*Y).$$

In order to compute its length, a curve in S^2 is therefore viewed as a curve in \mathbb{R}^3 . Consider for example the curve $\beta(t) = (0, \sin \pi t, \cos \pi t)$ for $t \in [0, 1]$, which follows a great circle between the North and South poles. We have

$$h_*(\dot{\beta}) = \pi \cos(\pi t) \frac{\partial}{\partial u_2} - \pi \sin(\pi t) \frac{\partial}{\partial u_3}.$$

Therefore, $\|\dot{\beta}(t)\| = h_*[\dot{\beta}(t)] = \pi$ and $L(\beta) = \pi$. This is actually the minimum length for curves in S^2 joining $(0, 0, 1)$ and $(0, 0, -1)$.

We shall return to these examples later. In the meantime, we would like to show that curves such as β , for which

$$L_t^s(\beta) = d(\beta(t), \beta(s))$$

are geodesics as defined in Section 2. We need to obtain a linear connection ∇ from the Riemannian metric, hopefully with some geometrically meaningful properties. Being motivated by considerations in \mathbb{R}^n , we are now torn between defining a geodesic as a ‘self-parallel’ curve (i.e., $\nabla_{\dot{\alpha}}\dot{\alpha} = 0$) and as a curve which minimizes distance. Since the self-parallel condition is a local one, using this definition of a geodesic eliminates some problems which otherwise arise. Let us illustrate one problem. If the domain of β is (as defined in example 3 above) extended to $[0, 2]$, the image of β is a great circle on S^2 and the total length of β is 2π . Obviously, 2π is not the distance from $\beta(0)$ to $\beta(2)$ since $\beta(0) = \beta(2)$. But the equation above is still valid for small $(s-t)$. The (global) length-minimizing property of β is lost when its domain is extended. To solve this problem, we shall find a connection so that a curve with a self-parallel tangent is *locally* length-minimizing.

NOTATION. For X and Y in $\mathfrak{X}(M)$, the C^∞ vector field $[X, Y]$, called the **Lie bracket** of X and Y , is defined by

$$[X, Y]_m f = X_m(Yf) - Y_m(Xf)$$

for all $f \in C^\infty(M)$. This makes sense because Yf and Xf are C^∞ functions.

To define ∇ , we shall specify the inner product of $\nabla_{X_m} Y$ and Z_m for all vector fields X, Y , and Z at all points $m \in M$. Define the **Levi-Civita or Riemannian connection** ∇ by

$$(3.1) \quad 2g_m(\nabla_{X_m} Y, Z_m) = X_m g(Y, Z) + Y_m g(X, Z) - Z_m g(X, Y) \\ + g_m([X, Y]_m, Z_m) + g_m([Z, X]_m, Y_m) + g_m([Z, Y]_m, X_m)$$

for all X, Y , and Z in $\mathfrak{X}(M)$. (Recall that $g(Y, Z)$, etc., are functions on M .) In a neighborhood U with coordinate functions x_1, \dots, x_n , let X_i denote the vector field $\partial/\partial x_i$. Then $[X_i, X_j] = 0$ because “mixed partial derivatives are equal” and (3.1) simplifies to

$$2g(\nabla_{X_i} X_j, X_k) = X_i g(X_j, X_k) + X_j g(X_k, X_i) - X_k g(X_i, X_j)$$

in U . A computation shows that ∇ satisfies the conditions for a linear connection.

This choice of ∇ is partially justified by the following theorem, known as the Fundamental Lemma of Riemannian Geometry (for a proof, see [5], p. 71).

THEOREM 3.1. *The connection ∇ defined above is the unique connection on M satisfying*

$$(3.2) \quad Xg(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z), \quad \text{and}$$

(3.3) *the torsion $T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y]$ is zero for all X, Y , and Z in $\mathfrak{X}(M)$.*

A linear connection satisfying (3.2) is called a **metric connection**. The equation expresses the directional derivative of the metric in terms of covariant derivatives, but it has more significance. Theorem 2.3 showed that the parallel translation operator is a critical part of the geometry of a connection. The operator most compatible with a Riemannian metric is an isometry [i.e., a map $\tau: T_m M \rightarrow T_q M$ such that $g_q(\tau X_m, \tau Y_m) = g_m(X_m, Y_m)$] because a Riemannian metric is nothing more than a point-wise inner product. Therefore a “natural” linear connection on a Riemannian manifold should have the property that parallel translation is an isometry. The following proposition tells us that this is indeed the case for a metric connection.

PROPOSITION 3.2. *Parallel translation is an isometry if and only if (3.2) holds for all vector fields X, Y , and Z .*

Proof. Assume (3.2) holds and let α be any curve in M . For Y and Z in $T_{\alpha(0)} M$, let Y_t and Z_t denote their parallel translates, $\tau_{\alpha(t)} Y$ and $\tau_{\alpha(t)} Z$ respectively, to $\alpha(t)$. Since Y and Z are parallel along α , if ∇ satisfies (3.2), then $\dot{\alpha}(t)g(Y, Z) = 0$. Hence

$$0 = \alpha_* \left(\frac{d}{dt} \right) g(Y, Z) = \frac{d}{dt} g_{\alpha(t)}(Y_t, Z_t)$$

and so $g_{\alpha(t)}(Y_t, Z_t)$ is constant, in other words

$$g_{\alpha(t)}(Y_t, Z_t) = g_{\alpha(0)}(Y, Z),$$

which is precisely the statement that $\tau_{\alpha(t)}$ is an isometry.

Assuming that $\tau_{\alpha(t)}$ is an isometry for any curve α , let X_m be in $T_m M$. To verify

(3.2) at m , let α be any curve with $\alpha(0) = m$ and $\dot{\alpha}(0) = X_m$. Both sides of (3.2) depend on Y and Z only along α . First consider vector fields Y and Z which are parallel along α . Then at m the left-hand side of (3.2) is zero since $g(Y, Z)$ is constant along α . The right-hand side is zero since $\nabla_{X_m} Y = \nabla_{X_m} Z = 0$. We have therefore verified (3.2) for this case. Now let $\{X_i\}$ be an orthonormal basis for $T_m M$ and let $\{X_i(t)\}$ denote its parallel translation to $\alpha(t)$. Since parallel translation is an isometry, $\{X_i(t)\}$ is an orthonormal basis for $T_{\alpha(t)} M$. Arbitrary differentiable vector fields Y and Z along α can be expressed $Y_{\alpha(t)} = \sum_i a_i(t) X_i(t)$ and $Z_{\alpha(t)} = \sum_i b_i(t) X_i(t)$, where the a 's and b 's are differentiable. Then

$$\begin{aligned} X_m g(Y, Z) &= \alpha_* \left(\frac{d}{dt} \Big|_0 \right) g(Y, Z) \\ &= \frac{d}{dt} \Big|_0 g_{\alpha(t)}(Y_{\alpha(t)}, Z_{\alpha(t)}) \\ &= \frac{d}{dt} \Big|_0 \sum_i a_i(t) b_i(t) \end{aligned}$$

and

$$\begin{aligned} g_m(\nabla_{X_m} Y, Z_m) + g_m(Y_m, \nabla_{X_m} Z) &= g_m \left(\sum_i \left[a_i(0) \nabla_{X_m} X_i + \frac{da_i}{dt} \Big|_0 X_i \right], \sum_j b_j(0) X_j \right) \\ &\quad + g_m \left(\sum_i a_i(0) X_i, \sum_j \left[b_j(0) \nabla_{X_m} X_j + \frac{db_j}{dt} \Big|_0 X_j \right] \right) \\ &= g_m \left(\sum_i \frac{da_i}{dt} \Big|_0 X_i, \sum_j b_j(0) X_j \right) + g_m \left(\sum_i a_i(0) X_i, \sum_j \frac{db_j}{dt} \Big|_0 X_j \right) \end{aligned}$$

because $\nabla_{X_m} X_i = 0$. Since $\{X_i\}$ is orthonormal, this is equal to

$$\sum_i \frac{da_i}{dt} \Big|_0 b_i(0) + a_i(0) \frac{db_i}{dt} \Big|_0$$

which is $X_m g(Y, Z)$. \parallel

A connection satisfying (3.3) is, locally, as close as possible to the standard Euclidean connection by the following theorem (see [10], Vol. II, p. 5-18 for a proof).

PROPOSITION 3.3. *The torsion of a connection ∇ is zero if and only if for every $m \in M$, there is a coordinate system $\{x_i\}$ for M near m so that*

$$\nabla_{\partial/\partial x_i} \frac{\partial}{\partial x_j} = 0,$$

or in classical notation, $\Gamma_{ij}^k = 0$.

A connection satisfying (3.3) is said to be **torsion-free**. Geodesics are determined

by a connection, as in Section 2, and metric connections are determined by their torsion. With our interest in geodesics, we note that two different metric connections may have the same geodesics (see [1], page 131, or [7]). By a **geodesic** of a Riemannian manifold we shall mean a geodesic of the Levi-Civita connection. Notice that differentiating $g(\dot{\alpha}(t), \dot{\alpha}(t))$ with respect to t shows that the tangent vectors to a geodesic α have constant length.

We come finally to the role of lengths of curves.

THEOREM 3.4. *For every point m in a Riemannian manifold M , there is a positive number λ and a neighborhood $U = \{q \in M \mid d(m, q) < \lambda\}$ such that*

(i) *any points q and r in U may be joined by a unique geodesic α whose image lies in U (unique up to a linear reparametrization),*

(ii) *the geodesic α is the unique curve joining q to r which has length $d(q, r)$ and*

(iii) *for each $q \in U$, there is a local coordinate system about q in which all geodesics α in U with $\alpha(0) = q$ have the form $\alpha(t) = (a_1 t, \dots, a_n t)$ with a_1, \dots, a_n constant.*

A proof may be found in [6], Vol. I, p. 166. This theorem says in particular that geodesics connecting nearby points have minimal length among all curves with the same endpoints. Furthermore, any point may be joined to any nearby point by a unique, length-minimizing geodesic. Locally the situation is similar to \mathbb{R}^n ; globally, however, this is not the case.

Given m and q , arbitrary points in a Riemannian manifold, there is not necessarily a curve of length $d(m, q)$ which joins m and q . For example, in $\mathbb{R}^2 - \{(0, 0)\}$ with the usual Riemannian metric as an open subset of \mathbb{R}^2 , every curve from $m = (-1, 0)$ to $q = (1, 0)$ has length at least 2 and for any $\varepsilon > 0$, there is a curve joining m and q with length less than $2 + \varepsilon$, hence $d(m, q) = 2$. However, there is no such curve of length 2. This phenomenon is due to the fact that $\mathbb{R}^2 - \{(0, 0)\}$ is not a complete metric space, as we shall see.

A manifold M with a linear connection ∇ is called **geodesically complete** if every geodesic $\alpha: [a, b] \rightarrow M$ determined by ∇ can be extended to a geodesic with domain \mathbb{R} .

THEOREM 3.5. (Hopf-Rinow). *A Riemannian manifold (with the Levi-Civita connection) is geodesically complete if and only if it is complete as a metric space, in which case any two points m and q in M may be joined by a curve of length $d(m, q)$.*

A proof may be found in [10], Vol. I, p. 9–55. The curve in Theorem 3.5 is actually a geodesic, as the next theorem shows.

THEOREM 3.6. *If α is a curve whose length realizes the distance between its endpoints, then α is a geodesic.*

Proof. Such a curve α must realize the distance between any two points in its image, or else a shorter (perhaps only piecewise C^∞) curve with the same endpoints could be found by replacing a section of α . Theorem 3.4 now guarantees that α is a geodesic in a neighborhood of every point in its image, therefore α is a geodesic. \parallel

The Christoffel symbols for the Levi-Civita connection are given in terms of the metric by $\Gamma_{ij}^k = g(\nabla_{X_i} X_j, X_k)$, where $\{X_i = \partial/\partial x_i\}$ is an orthonormal basis for $T_m M$. For actual computations, it is convenient to write the metric as a matrix-valued function on coordinate neighborhoods $[g_{ij}] = [g(X_i, X_j)]$ with inverse g^{-1} . The Christoffel symbols may then be computed:

$$(3.4) \quad \Gamma_{ij}^k = \frac{1}{2} \sum_h g_{hk}^{-1} \left(\frac{\partial g_{ih}}{\partial x_j} + \frac{\partial g_{jh}}{\partial x_i} - \frac{\partial g_{ij}}{\partial x_h} \right).$$

Let us revisit the examples of this section. All complete Riemannian manifolds, they represent the classical Euclidean, hyperbolic, and spherical geometries.

1. In \mathbb{R}^n with its standard metric, differentiating the constant function $g(\partial/\partial u_i, \partial/\partial u_j)$ with respect to $\partial/\partial u_k$ shows that $\nabla_{\partial/\partial u_i} \partial/\partial u_j = 0$ for all i and j . The Levi-Civita connection of the standard metric is therefore the standard connection. The fact that $\Gamma_{ij}^k = 0$ for all i, j , and k was known to Christoffel. In this geometry, two distinct points lie on a unique line, and given a line and a point not on it, there is a unique line through the point which does not intersect the given line. By "line" we mean the *points* in the image of a geodesic defined on \mathbb{R} .

2. *The Poincaré upper half-plane.* Let $H \subset \mathbb{R}^2$ be the upper half-plane with the metric given as before by $g_{ij} = \delta_{ij}/u_2^2$. Then $g_{11}^{-1} = g_{22}^{-1} = u_2^2$ and $g_{12}^{-1} = g_{21}^{-1} = 0$. We can easily compute the Christoffel symbols for H using Equation 3.4: the only non-zero ones are $\Gamma_{12}^1 = \Gamma_{21}^1 = -\Gamma_{11}^2 = \Gamma_{22}^2 = -\frac{1}{u_2^2}$. The equations for a geodesic $\alpha = (\alpha_1, \alpha_2)$, where $\alpha_2 > 0$, now become

$$\frac{d^2 \alpha_1}{dt^2} - \frac{2}{\alpha_2(t)} \frac{d\alpha_1}{dt} \frac{d\alpha_2}{dt} = 0$$

and

$$\frac{d^2 \alpha_2}{dt^2} + \frac{1}{\alpha_2(t)} \left(\frac{d\alpha_1}{dt} \right)^2 - \frac{1}{\alpha_2(t)} \left(\frac{d\alpha_2}{dt} \right)^2 = 0.$$

The only solutions to these equations are

$$\alpha_1(t) = a + b \tanh(rt + c), \alpha_2(t) = b \operatorname{sech}(rt + c)$$

and

$$\alpha_1(t) = a', \alpha_2(t) = b' e^{r't+c'}$$

where the a 's, b 's, c 's, and r 's are constant. This manifold is also called the hyper-

bolic upper half-plane. A “line” (point-set image $\alpha(\mathbb{R})$ of a geodesic α) is either a semi-circle centered on the horizontal axis or a vertical straight line (Fig. 3), and two points determine a unique line. Notice that if “parallel” means non-intersecting, in this geometry there are an infinite number of lines through a given point (a, b) and parallel to a given line not containing (a, b) . Thus Euclid’s Fifth Postulate is not satisfied.

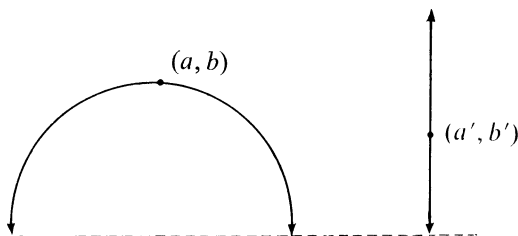


FIG. 3

3. *The Riemann sphere.* Let S^2 possess the metric induced by \mathbb{R}^3 . It can be shown that a tangent vector to S^2 at m is precisely a tangent vector to \mathbb{R}^3 at m which is orthogonal to m (as a vector in \mathbb{R}^3), and therefore a curve α in S^2 is a geodesic in S^2 only if the Euclidean derivative $\nabla_{\dot{\alpha}}\dot{\alpha}$ (as a vector field along α in \mathbb{R}^3) is everywhere perpendicular to S^2 . It can be shown that geodesics in S^2 are precisely the constant-speed parametrizations of great circles. This manifold is also called spherical space. Notice that antipodal points lie on infinitely many distinct “lines” (with the same interpretation as in Example 2), while any other pair of points lie on a unique line. Since two distinct lines intersect in exactly two points, there are no parallel lines, and again Euclid’s Fifth Postulate does not hold.

4. *Geometry as a connection of a fibre bundle.* We saw in Section 2 that the geometric concept of covariant derivative leads naturally to the equally geometric concept of parallel translation of tangent vectors along a curve, i.e., for each curve α in M , an isomorphism $\tau_{\alpha(t)}: T_{\alpha(0)}M \rightarrow T_{\alpha(t)}M$. We shall now interpret τ_{α} as a map from a fiber bundle to itself which satisfies certain conditions.

Let M be an n -manifold. A **frame** on M is a point m of M (called the **origin** of the frame) together with a basis for T_mM . Let $L(M)$ be the collection of all frames on M :

$$L(M) = \{(m, X_1, \dots, X_n) \mid m \in M \text{ and } \{X_i\} \text{ is a basis for } T_mM\}.$$

$L(M)$ is called the **frame bundle** of M . Let $\pi: L(M) \rightarrow M$ be given by $\pi[(m, X_1, \dots, X_n)] = m$. We shall make $L(M)$ into a manifold as follows: Let $(U_{\gamma}, \phi_{\gamma})$ be a chart for M and let $V_{\gamma} = \pi^{-1}(U_{\gamma})$, i.e.,

$$V_{\gamma} = \{(m, X_1, \dots, X_n) \in L(M) \mid m \in U_{\gamma}\}.$$

For any point m in U_{γ} ,

$$\left\{ \frac{\partial}{\partial x_1} \Big|_m, \dots, \frac{\partial}{\partial x_n} \Big|_m \right\} \text{ is a basis for } T_m M,$$

(where we have suppressed the subscript γ by writing x_i for $u_i \circ \phi_\gamma$). Since any two bases for a given vector space (in our case, $T_m M$) differ by a non-singular matrix, given any frame (m, X_1, \dots, X_n) , there is a matrix $A = [a_{ij}]$ in $Gl(n, \mathbb{R})$ such that A carries $\{\partial/\partial x_i|_m\}$ to $\{X_i\}$, i.e.,

$$X_i = \sum_j a_{ij} \frac{\partial}{\partial x_j} \Big|_m.$$

Writing G for $Gl(n, \mathbb{R})$, we have maps $\psi_\gamma: V_\gamma \rightarrow U_\gamma \times G$ defined by $\psi_\gamma[(m, X_1, \dots, X_n)] = (m, A^t)$. We may put a topology on $L(M)$ which makes each ψ_γ continuous by choosing as a base $\{\psi_\gamma^{-1}(W) \mid W \text{ is open in } U_\gamma \times G\}$. We now define a chart for $L(M)$ to be a pair $(V_\gamma, \tilde{\phi}_\gamma)$, where $\tilde{\phi}_\gamma: V_\gamma \rightarrow \mathbb{R}^{n+n^2}$ is given by

$$\tilde{\phi}_\gamma[(m, X_1, \dots, X_n)] = (x_1(m), \dots, x_n(m), a_{11}, a_{12}, \dots, a_{nn}).$$

It is easy to see that if $V_\gamma \cap V_\delta \neq \emptyset$,

$$\begin{aligned} & \tilde{\phi}_\delta \circ \tilde{\phi}_\gamma^{-1}(x_1, \dots, x_n, a_{11}, a_{12}, \dots, a_{nn}) \\ &= (\phi_\delta \circ \phi_\gamma^{-1}(x_1, \dots, x_n), J_{\phi_\delta \circ \phi_\gamma^{-1}}^t(a_{11}, \dots, a_{1n}), \dots, J_{\phi_\delta \circ \phi_\gamma^{-1}}^t(a_{n1}, \dots, a_{nn})) \end{aligned}$$

and so the compatibility condition is satisfied. Here $J_{\phi_\delta \circ \phi_\gamma^{-1}}^t$ is the transpose of the $n \times n$ Jacobian matrix $J_{\phi_\delta \circ \phi_\gamma^{-1}}$. With the given topology and charts, $L(M)$ is a differentiable manifold of dimension $n + n^2$. $L(M)$ is sometimes called the bundle of bases of M .

We also have an "action" of G on $L(M)$, specifically the map $\Phi: L(M) \times G \rightarrow L(M)$ given by

$$\Phi((m, X_1, \dots, X_n), A) = (m, \sum_j a_{j1} X_j, \dots, \sum_j a_{jn} X_j),$$

where $A = [a_{ij}]$ as usual. We will write b^A for $\Phi(b, A)$ if $b \in L(M)$. Observe that for all $b \in L(M)$, $(b^A)^B = b^{AB}$ for all A and B in G , and $b^A = b$ if and only if A is the identity (matrix) of G .

A curve $\tilde{\alpha}: I \rightarrow L(M)$ is called a **lift** of α ($\alpha: I \rightarrow M$) if $\pi \circ \tilde{\alpha} = \alpha$. For $b \in \pi^{-1}(\alpha(0))$, $\tilde{\alpha}$ is called a **lift at b** if in addition $\tilde{\alpha}(0) = b$. The geometric importance of the frame bundle lies in the following theorem about lifts.

THEOREM 4.1. *Assigning a parallel translation τ_α along each curve α in M is equivalent to assigning to each curve α and each $b \in \pi^{-1}(\alpha(0))$, a unique lift $\tilde{\alpha}_b$ of α at b such that*

$$\tilde{\alpha}_b(t) = [\tilde{\alpha}_b(t)]^A$$

for all t in the domain of α .

(The condition above means that the lift of α at b^A is the lift of α at b acted on by A at each point. We call this property **equivariance**. For reasons which will appear later, the equivariant lift in the theorem is called the **horizontal lift** of α at b .)

Proof. Letting $b = (\alpha(0), X_1, \dots, X_n) \in \pi^{-1}(\alpha(0))$, the correspondence is given by

$$\begin{aligned} \tau_{\alpha(t)} \left(\sum_i c_i X_i \right) &= \sum_i c_i Y_i(\alpha(t)) \leftrightarrow \\ \tilde{\alpha}_b(t) &= (\alpha(t), Y_1(\alpha(t)), \dots, Y_n(\alpha(t))). \end{aligned}$$

Assuming that each curve α has a unique horizontal lift $\tilde{\alpha}_b$ at b , $\tilde{\alpha}_b(t)$ has the form on the right, where $\{Y_i(\alpha(t))\}$ is a basis for $T_{\alpha(t)}M$. We then define τ_α by the expression on the left. It is easy to check that τ_α is independent of the choice of b (precisely because of equivariance) and that τ_α is an isomorphism.

Conversely, given τ_α for each curve α , $\tilde{\alpha}$ is defined by the right-hand side. Certainly $\tilde{\alpha}_b$ is a lift of α to b since $\tau_{\alpha(0)}$ is the identity on $T_{\alpha(0)}M$. The equivariance follows from Theorem 2.4. \parallel

In this proof, we obtained $\tilde{\alpha}_b$ by parallel translation along α of the basis represented by b . One may picture this (for 2-manifolds) as in Fig 4a. After superimposing the picture of $\tilde{\alpha}_b$ on M (Fig. 4b), we see that $\tilde{\alpha}_b$ looks very much like a moving frame, where the dotted vectors are obtained by parallel translation.

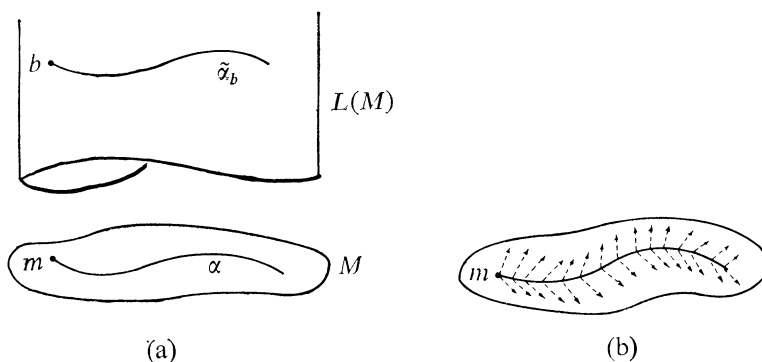


FIG. 4

Let us recap briefly the situation. We have two manifolds, M and $L(M)$, a Lie group G ($Gl(n, \mathbb{R})$) acting on $L(M)$, and a map $\pi: L(M) \rightarrow M$. Our theorem says that the geometric structure on M of parallel translation is precisely the same as unique equivariant path-lifting of curves in M to curves in $L(M)$ with specified initial points. This last concept is analogous to that of unique path-lifting to covering spaces (with a covering space of M replacing $L(M)$ and the group of covering trans-

formations replacing G). Indeed, the problem of the existence of a lift $\tilde{\alpha}$ is really a special case of the “lifting problem” in topology. The notion and the ability to lift curves will be seen to be topological concepts, whereas the uniqueness of the lift will be a geometric concept. Why can we lift curves in M to curves in $L(M)$? It is because $\pi^{-1}(U_\gamma)$ is homeomorphic (via ψ_γ) to $U_\gamma \times G$. If α is any curve in U_γ and h is any function from U_γ into G , then $\tilde{\alpha}(t) = \psi_\gamma^{-1}(\alpha(t), h \circ \alpha(t))$ is a lift (not necessarily the horizontal lift) of α . It is this “local product” structure together with the group action which play a central role in our development of geometry. We shall define a principal fiber bundle by imitating the essential features of the frame bundle, and be able to put a “geometry” on it,

Let P be a manifold. We say that a Lie group G **acts freely** on P if there is a map from $P \times G$ into P (we will write $(p, g) \mapsto p^g$) such that for all $p \in P$, $(p^g)^h = p^{gh}$ for all g and h in G , and $p^g = p$ for any $p \in P$ if and only if g is the identity in G . Points p and q in P are called **equivalent** under (the action of) G if $q = p^g$ for some $g \in G$.

A (differentiable) **principal fiber bundle** $G \rightarrow P \xrightarrow{\pi} M$ consists of manifolds P (the bundle space or total space) and M (the base space), a Lie group G (the structural group), and a C^∞ map $\pi: P \rightarrow M$ such that

- (i) G acts freely on P ,
- (ii) M is the quotient space of P under the action of G , so that $\pi(p) = \pi(q)$ if and only if p and q are equivalent under G , and
- (iii) for every U_γ in some open covering of M , there is a diffeomorphism $\psi_\gamma: \pi^{-1}(U_\gamma) \rightarrow U_\gamma \times G$ which commutes with the action of G , i.e., if $\psi_\gamma(p) = (\pi(p), h)$, then $\psi_\gamma(p^g) = (\pi(p), hg)$.

For $m \in M$, the “fiber” $\pi^{-1}(m)$ is diffeomorphic to G , by condition (ii). Observe that P is locally the product of M and G but that in general, P need not be diffeomorphic to $M \times G$. (In particular, $L(M) \neq M \times Gl(n, \mathbb{R})$.) We shall mention only three more examples of principal fiber bundles; others are presented, for example, in [11].

Examples: 1. (The trivial bundle.) Let $P = M \times G$ and π be the projection onto the first factor. Then G acts on P by $((m, g), h) \rightarrow (m, gh)$.

2. Let P be any covering space for M , $\pi: P \rightarrow M$ be the covering map, and G be the group of deck transformations (with the discrete topology).

3. (The canonical \mathbb{C}^* bundle over $\mathbb{C}P^n$.) Let $P = \mathbb{C}^{n+1} - \{0\}$ (complex $(n+1)$ -space minus the origin) and $G = \mathbb{C}^*$ (the non-zero complex numbers). Complex projective n -space $\mathbb{C}P^n$ is defined as follows: we say that for z_1 and z_2 in P , $z_1 \sim z_2$ if there is $\lambda \in G$ such that $z_1 = \lambda z_2$. Then the set $\mathbb{C}P^n$ of \sim -equivalence classes of P is a $2n$ -manifold and $G \rightarrow P \rightarrow \mathbb{C}P^n$ is a principal fiber bundle.

If p is a point in a bundle space, the set $V_p = \{X \in T_p P \mid \pi_*(X) = 0\}$ is called the vertical subspace at p . It is clear that a fiber $\pi^{-1}(m)$ is a manifold which sits

inside P and whose tangent space at each point p is the vertical subspace V_p (Fig. 5), for $\pi_*(X) = 0$ if and only if Xf depends only on the restriction of f to $\pi^{-1}(m)$.

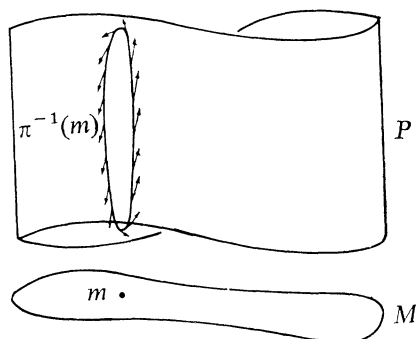


FIG. 5

It is a well-known result in algebraic topology that any path in the base of a fiber bundle may be lifted to the total space (because of the local product structure). The lifting at a specified initial point is unique if and only if the fiber has no non-constant paths (i.e., the fiber is discrete). That lifts are not in general unique is not surprising from our viewpoint because, in Riemann's language, the fiber bundle is the "space" and something (a method of unique path-lifting) must be added to provide the geometry. The ambiguity of the lift is that the lifted curve could try to "travel up the fiber" as well as move "horizontally." (Compare with covering spaces, where the lift cannot go up.) We shall define a connection on a principal fiber bundle, which is nothing more than a direction to go, or many directions not to go.

Let ξ denote a principal fiber bundle $G \rightarrow P \rightarrow M$ and n be the dimension of M . A **connection** H on ξ is an assignment for each $p \in P$ of an n -dimensional subspace $H_p \subset T_p P$, called the **horizontal subspace** at p , such that for each p ,

- (i) $T_p P = V_p \oplus H_p$,
- (ii) $(R_g)_*(H_p) = H_{(pg)}$ for all $g \in G$, and
- (iii) If $h: T_p P \rightarrow H_p$ is projection and X is a vector field on P , then hX is also a vector field on P . (This is a differentiability condition on H .)

We have written R_g for the diffeomorphism of P given by $p \mapsto p^g$.

THEOREM 4.2. *Giving a connection on ξ is equivalent to giving unique equivariant path-lifting in ξ .*

Proof (sketch). Assume that a connection H is given on ξ and let α be a simple curve in M . (For a proof without this assumption on α , see [1], pp. 77–79.) Since $\pi_* X = 0$ if and only if X is vertical, we see that the restriction of π_* to any horizontal subspace H_p is one-to-one, hence an isomorphism (by dimensions) of H_p onto $T_p M$. For each $p \in \pi^{-1}(\alpha(t))$, let X_p be the unique horizontal vector at p

which projects onto $\dot{\alpha}(t)$. These vectors may be extended to a vector field X on P which at each point lies in the horizontal subspace. Now for $p \in \pi^{-1}(\alpha(0))$, let $\tilde{\alpha}_p$ be the (unique) integral curve of X such that $\tilde{\alpha}(0) = p$ (Fig. 6). The compactness of I insures that $\tilde{\alpha}_p$ may be defined on all of I , if necessary by piecing together curves defined on small intervals. The tangent vectors to $\tilde{\alpha}_p$ are clearly horizontal and $\tilde{\alpha}_p$ is a C^∞ curve which projects onto α (this is not clear!). The equivariance follows from our construction and the second condition for H .

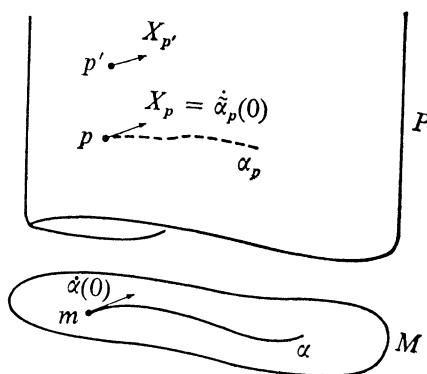


FIG. 6

Now suppose we have unique equivariant path-lifting on ξ . To define H_p for $p \in P$, let $\alpha_1, \dots, \alpha_n$ be curves in M such that $\alpha_i(0) = m = \pi(p)$ and $\{\dot{\alpha}_i(0)\}$ forms a basis for $T_m M$. (The α_i 's may be taken to be integral curves of the vectors in any basis for $T_m M$.) Let $\tilde{\alpha}_i$ be the lift of α_i at p . Define H_p to be the span of $\{\dot{\tilde{\alpha}}_i(0)\}$ (see Fig. 7). A little work shows that H_p is independent of the choice of the α_i 's and that $p \mapsto H_p$ is a connection. \parallel

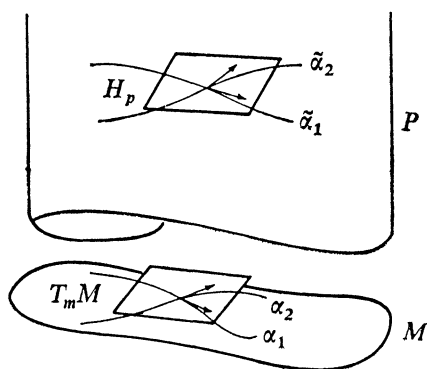


FIG. 7

The definition of connection in this setting is due to Ehresmann [3] although

it is implicit in some work of Cartan. The authors feel that even in this setting there is the dichotomy of Riemann—the concept of space (the principal fiber bundle) and the additional structure which yields the geometry (the connection), and that the geometric content of the massive structure of modern differential geometry is still apparent.

References

1. R. L. Bishop and R. J. Crittenden, *Geometry of Manifolds*, Academic Press, New York, 1964.
2. S. S. Chern, Some new viewpoints in differential geometry in the large, *Bull. Amer. Math. Soc.*, 52 (1946) 1–30.
3. C. Ehresmann, Les connexions infinitésimales dans un espace fibré différentiable, *Colloque de Topologie, Bruxelles* (1950) 29–55.
4. H. Flanders, Differential forms, MAA Studies No. 4 in *Global Geometry and Analysis*, edited by S. S. Chern, 1967.
5. N. J. Hicks, *Notes on Differential Geometry*, Van Nostrand, Princeton, N. J., 1965.
6. S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, Vol. I and II. Interscience, New York, 1963, and 1969.
7. R. S. Millman, Geodesics in metrical connections, *Proc. Amer. Math. Soc.*, 30 (1971) 551–555.
8. K. Nomizu, Recent developments in the theory of connections and holonomy groups, *Advances in Math.*, 1 (1961) 1–49.
9. D. E. Smith, *A Source Book in Mathematics*, Vol. II. Dover, New York, 1959.
10. M. Spivak, *Differential Geometry*, Vol. I and II. Published by Michael Spivak, 1970.
11. N. Steenrod, *The Topology of Fiber Bundles*, Princeton Univ. Press, 1951.

AN INTRODUCTION TO MATROID THEORY

R. J. WILSON, The Open University, England

1. Introduction. In this expository article, we shall be presenting a survey of some of the most important aspects of matroid theory, a branch of combinatorial mathematics which has come very much to the fore in the last few years.

The subject originates from a fundamental paper of Hassler Whitney [35], which appeared in 1935. He had just spent several years working in the field of graph theory, and had noticed several similarities between the ideas of independence and rank in graph theory and those of linear independence and dimension in the study of vector spaces. In his paper Whitney used the concept of a matroid to formalize these similarities. A **matroid** is essentially a set with some kind of ‘independence

Robin Wilson did his undergraduate work at Balliol College, Oxford, and his graduate work at the University of Pennsylvania and M.I.T. His Penn. Ph.D. on sieve methods was written under N.C. Ankeny. He was a Lecturer at Jesus College, Oxford, and now is Lecturer at the Open University. The present article is derived from his lectures at the Combinatorial Analysis Institute, Bowdoin College. He is the author of *Introduction to Graph Theory* (Oliver & Boyd, Edinburgh, 1972). *Editor.*

structure' defined on it; the name 'matroid' arose from his consideration of the independence of the columns of a matrix. At about the same time, B. L. van der Waerden [32] rediscovered the idea of a matroid while trying to formalize the definitions of linear and algebraic independence.

The work of Whitney and van der Waerden was largely ignored for over twenty years (with the important exceptions of a paper of S. MacLane [15] in 1936, and one of R. Rado [27] in 1942) until a breakthrough occurred in 1958 when W. T. Tutte characterized those matroids which arise from graphs (see [31]). Later, in 1965, J. Edmonds and D. R. Fulkerson [8] (and independently, L. Mirsky and H. Perfect [20] and Brualdi and Scrimger [4]) recognized the importance of matroids in transversal theory. Since then, a large number of combinatorialists have contributed to the subject, and there is already an impressive literature in the field.

In this article, no previous knowledge of graph theory or transversal theory is assumed. We shall develop the necessary background material for these subjects in Sections 2–4, and these sections are written in such a way as to motivate the material which follows. In Sections 5–7, several equivalent definitions of a matroid are presented, together with a wide variety of examples. After discussing matroid duality in some detail, we then show (Sections 10–12) how matroid theory can be used to simplify various ideas in graph theory and transversal theory. The article concludes with a brief discussion of some recent work in the subject. Our aim throughout is to show that matroid theory is far from being 'generalization for generalization's sake'; on the contrary, it gives us a deeper insight into several problems in transversal theory, as well as including among its applications simple proofs of results in graph theory which are awkward to prove by more traditional methods.

The treatment of the subject given here is somewhat similar to that of the last chapter of the author's recent introductory text on graph theory [36]; several proofs which have been omitted from this article may be found in this book. We shall be interested only in **finite matroids** (i.e., matroids defined on finite sets), and shall always use $|E|$ to denote the number of elements in a set E . The reader who is interested in matroids defined on infinite sets should refer to Rado [28] or Brualdi and Scrimger [4] for further details. Finally, we should like to point out that although due credit has been given where possible to those responsible for results cited in this paper, several of these results are so firmly embedded in the folk-lore of the subject, that to give such due credit has been impossible. We should consequently like to apologize in advance to anyone who feels that he (or she) has been overlooked.

2. Some results in Graph Theory. In this section, we present a brief survey of those results of graph theory which we shall need later. The reader is referred to [36] for a fuller treatment of the subject.

A **graph** G is defined to be a pair $(V(G), E(G))$, where $V(G)$ is a finite non-empty set of elements (called **vertices**), and $E(G)$ is a finite family of unordered pairs of elements of $V(G)$ (called **edges**). An example of a graph is given in figure 1; in this

example, $V(G)$ is the set $\{u, v, w, z\}$, and $E(G)$ consists of the edges $\{u, v\}$, $\{v, v\}$, $\{v, w\}$, $\{v, w\}$, $\{u, w\}$ and $\{w, z\}$. The edge $\{v, v\}$ is called a **loop**, and the two edges of the form $\{v, w\}$ are called **multiple edges**; any graph containing no loops or multiple edges is called a **simple graph**. The edge $\{u, v\}$ is said to **join** the vertices u and v , and u and v are then said to be **incident** to this edge. A graph, each of whose vertices belongs to $V(G)$ and each of whose edges belongs to $E(G)$, is called a **subgraph** of G . We shall call two graphs G and G' **isomorphic** if there is a one-one correspondence between their sets of vertices with the property that the number of edges joining any two vertices of G is equal to the number of edges joining the corresponding vertices of G' .

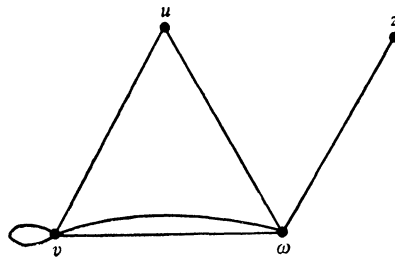


FIG. 1

If G is a graph, a **path** in G is a finite sequence of distinct edges of the form

$$\{v_0, v_1\}, \{v_1, v_2\}, \dots, \{v_{m-1}, v_m\}.$$

A graph G is called **connected** if, given any two vertices v and w , there is a path connecting v and w . Any graph which is not connected may be split up into a finite number of connected subgraphs, called **components**; for example, the graph in figure 2 has three components.

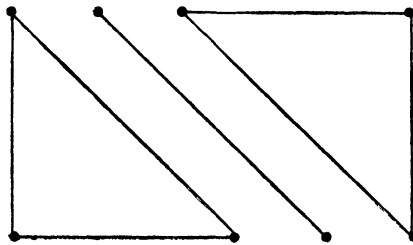


FIG. 2

A non-empty path in which all the vertices are distinct (except for v_0 and v_m , which are the same) is called a **circuit**; for example, any loop or any pair of multiple edges forms a circuit. A graph in which every circuit contains an even number of

edges is called a **bipartite graph** (see figure 3); note that in a bipartite graph, the set of vertices can be partitioned into two sets in such a way that each edge joins two vertices, one from each set. A **cutset** of a graph G is a set of edges whose removal increases the number of components of G , and which is minimal with respect to this property; for example, in figure 1, the edges $\{u, v\}$ and $\{u, w\}$ form a cutset. The following results can now be easily proved:

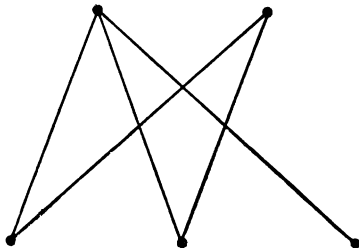


FIG. 3

(G1) (i) If C_1 and C_2 are distinct circuits of a graph G , each containing an edge e , then there exists a circuit in $C_1 \cup C_2$ which does not contain e ; similarly, (ii) if C_1^* and C_2^* are distinct cutsets of G , each containing an edge e , then there exists a cutset in $C_1^* \cup C_2^*$ which does not contain e .

(G2) If C is a circuit of a graph G and C^* is a cutset, then the number of edges of G common to C and C^* is even.

A graph which contains no circuits is called a **forest**, and a connected forest is a **tree**. If G is a connected graph, then a **spanning tree** T of G is a tree which contains every vertex of G and all of whose edges are edges of G (see figure 4). Similarly, if G is any graph, we define a **spanning forest** of G to be a forest obtained by taking a spanning tree of each component of G . The following results are now easily proved:

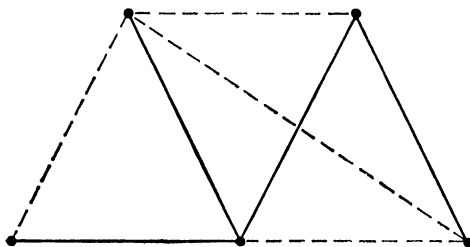


FIG. 4

(G3) No spanning forest of G contains another spanning forest as a proper subgraph.

(G4) If T_1 and T_2 are two spanning forests of G , and e is an edge of T_1 , then there exists an edge f of T_2 with the property that $(T_1 - \{e\}) \cup \{f\}$ (the graph obtained from T_1 on replacing e by f) is also a spanning forest of G .

By repeatedly using these two results, (G3) and (G4), one can easily deduce the following:

(G5) Any two spanning forests of G contain the same number of edges.

If G contains n vertices and k components, then the number of edges in any spanning forest is $n - k$; this number is called the **rank** of G and is denoted by $\kappa(G)$. The number of edges which must be removed from G to produce a spanning forest will be denoted by $\gamma(G)$, and is equal to $m - n + k$, where m denotes the number of edges of G . Some of the most important properties of the rank function are described in the following proposition:

(G6) The rank function κ satisfies the following properties: (i) for each subgraph H of G , $0 \leq \kappa(H) \leq |E(H)|$; (ii) if H is a subgraph of K , then $\kappa(H) \leq \kappa(K)$; (iii) for any subgraphs H and K of G , $\kappa(H \cup K) + \kappa(H \cap K) \leq \kappa(H) + \kappa(K)$.

We conclude this section with two simple results which combine the concepts of circuit and cutset with that of a spanning forest:

(G7) (i) Every cutset of a graph G has an edge in common with every spanning forest; (ii) every circuit of G has an edge in common with the complement of any spanning forest (i.e., the graph obtained by removing from G the edges of the spanning forest).

3. Some results on vector spaces and projective spaces. If V is a finite-dimensional vector space over a field F , a **basis** of V is a linearly independent set of elements which span V . It is well known that the bases of V satisfy the following properties:

(V1) No basis of V properly contains another basis of V .

(V2) If B_1 and B_2 are bases of V , and v is an element of B_1 , then there exists an element w of B_2 with the property that $(B_1 - \{v\}) \cup \{w\}$ is also a basis of V .

It is now easy to prove the following:

(V3) Any two bases of V contain the same number of elements.

As in the previous section, we can obtain a **rank function** by defining the **rank** $r(A)$ of each subset A of V to be the dimension of the subspace of V spanned by the vectors in A . The following proposition is then easily proved:

(V4) The rank function r satisfies the following properties: (i) for each subset

A of V , $0 \leq r(A) \leq |A|$; (ii) if $A \subseteq B$, then $r(A) \leq r(B)$; (iii) for any A and B , $r(A \cup B) + r(A \cap B) \leq r(A) + r(B)$.

A result analogous to (V4) holds also for projective spaces. If P is a projective space of dimension n , then we define the **rank** $r(A)$ of a finite set A of elements of P to be *one more than* the dimension of the projective subspace spanned by A . Clearly, this rank function also satisfies the properties (i), (ii) and (iii) of (V4).

4. Some results in transversal theory. In this section, we present a brief introduction to transversal theory. A fuller treatment of this subject, including proofs of the results stated here, may be found in Mirsky's survey article [19] or in his book [18].

Let E be a finite set, and $\mathcal{S} = (S_1, \dots, S_m)$ be a family of non-empty subsets of E . A **transversal** (or **system of distinct representatives**) of \mathcal{S} is a set of m distinct elements of E , one chosen from each of the subsets S_i ; a **partial transversal** of \mathcal{S} is a transversal of some subfamily of \mathcal{S} . For example, the family $\mathcal{S} = (S_1, S_2, S_3)$ of subsets of $E = \{a, b, c, d\}$, where $S_1 = \{b, c, d\}$, $S_2 = S_3 = \{a\}$, has no transversal, but its partial transversals are easily seen to be \emptyset , $\{a\}$, $\{b\}$, $\{c\}$, $\{d\}$, $\{a, b\}$, $\{a, c\}$ and $\{a, d\}$. Note that the situation can be represented by a bipartite graph in which each edge joins one of the S_i to one of its elements (see figure 5); a partial transversal then corresponds to a set of edges, no two of which have a vertex in common.

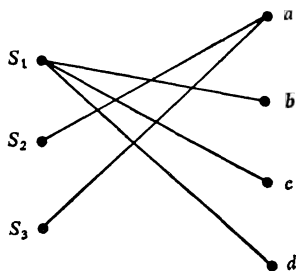


FIG. 5

In 1935, Philip Hall obtained necessary and sufficient conditions for the family \mathcal{S} to have a transversal:

(T1) (Hall's 'marriage' theorem): \mathcal{S} has a transversal if and only if, for each k satisfying $1 \leq k \leq |E|$, the union of any k of the subsets S_i contains at least k elements.

By a straightforward (but somewhat technical) argument, one can also prove the following result:

(T2) \mathcal{S} has a transversal containing a given subset A if and only if (i) \mathcal{S} has a transversal, and (ii) A is a partial transversal of \mathcal{S} .

The reason for introducing transversal theory should become clearer after we have stated the next two propositions. In the following, a **maximal partial transversal** of \mathcal{S} is a partial transversal of \mathcal{S} which is not properly contained in any other partial transversal; it follows that if \mathcal{S} actually has a transversal, then every maximal partial transversal of \mathcal{S} is a transversal.

(T3) *No maximal partial transversal of \mathcal{S} properly contains another maximal partial transversal.*

(T4) *If T_1 and T_2 are maximal partial transversals of \mathcal{S} , and x is an element of T_1 , then there exists an element y of T_2 with the property that $(T_1 - \{x\}) \cup \{y\}$ is a maximal partial transversal of \mathcal{S} .*

By repeatedly using these two results, one can easily deduce the following:

(T5) *Any two maximal partial transversals of \mathcal{S} contain the same number of elements.*

As before, we can define a rank function σ on the set of subsets of E , by defining $\sigma(A)$ to be the number of elements in the largest partial transversal of \mathcal{S} contained in A ; we can now prove the following:

(T6) *The rank function σ satisfies the following properties: (i) for each subset A of E , $0 \leq \sigma(A) \leq |A|$; (ii) if $A \subseteq B$, then $\sigma(A) \leq \sigma(B)$; (iii) for any $A, B \subseteq E$, $\sigma(A \cup B) + \sigma(A \cap B) \leq \sigma(A) + \sigma(B)$.*

The reader should note the similarities between (G3), (V1) and (T3), between (G4), (V2) and (T4), between (G5), (V3) and (T5), and between (G6), (V4) and (T6).

5. The definition of a matroid. Motivated by the results of the three previous sections, we now give several different definitions of a matroid; proofs of their equivalence may be found in Whitney's original paper [35]. The reader who finds this section heavy-going may wish to refer forward to Section 6, where several examples are given.

A **matroid** $M = (E, \mathcal{B})$ consists of a non-empty finite set E , together with a non-empty collection \mathcal{B} of subsets of E (called **bases**) satisfying the following properties:

(B*i*) *No base properly contains another base.*

(B*ii*) *If B_1 and B_2 are bases, and x is an element of B_1 , then there exists an element y of B_2 with the property that $(B_1 - \{x\}) \cup \{y\}$ is also a base.*

As in the previous three sections, one can easily deduce from these two properties that *any two bases of a matroid contain the same number of elements* (sometimes called the **rank** of the matroid). Note that this result generalizes (G5), (V3) and (T5).

If M is a matroid on a set E , we say that a subset A of E is an **independent set** if A is contained in some base of M . It follows that the bases of M are precisely the maximal independent sets, and hence that the matroid is completely determined by listing its independent sets. It seems reasonable to expect, therefore, that there may

be a simple definition of a matroid in terms of its independent sets. One such definition is as follows:

A **matroid** $M = (E, \mathcal{I})$ consists of a non-empty finite set E , together with a non-empty collection \mathcal{I} of subsets of E (called **independent sets**) satisfying the following properties:

- (\mathcal{I} i) *any subset of an independent set is an independent set;*
- (\mathcal{I} ii) *if I and J are independent sets, and $|J| > |I|$, then there exists an element x belonging to J but not to I with the property that $I \cup \{x\}$ is an independent set.*

Using this definition, it is an easy exercise to show that *any independent set can be extended to a base*, and that *if A is a subset of E , then any two maximal independent subsets of A contain the same number of elements*. (The reader should interpret these results in terms of graphs, vector spaces, and transversals.) We can also use independent sets to formulate the definition of the isomorphism of matroids: two matroids M_1 and M_2 are **isomorphic** if there is a one-one correspondence between their underlying sets E_1 and E_2 , with the property that a set of elements of E_1 is independent in M_1 if and only if the corresponding set of elements of E_2 is independent in M_2 .

Our third definition of a matroid is very similar to Whitney's original definition; this is given in terms of a rank function and generalizes the results (G6), (V4) and (T6) above.

If $M = (E, \mathcal{I})$ is a matroid defined in terms of its independent sets, then the **rank** $\rho(A)$ of a subset A of E is defined to be the number of elements in the largest independent set contained in A . It follows that a subset A of E is independent if and only if $\rho(A) = |A|$, and that if B is a base, then $|B| = \rho(B) = \rho(E)$. Since a matroid is completely determined by its rank function ρ , we can redefine a matroid in terms of it, as follows:

A **matroid** $M = (E, \rho)$ consists of a non-empty finite set E , together with an integer-valued function ρ (called its **rank function**) which is defined on the set of subsets of E and which satisfies the following properties:

- (ρ i) *for each subset A of E , $0 \leq \rho(A) \leq |A|$;*
- (ρ ii) *if $A \subseteq B \subseteq E$, then $\rho(A) \leq \rho(B)$;*
- (ρ iii) *for any $A, B \subseteq E$, $\rho(A \cup B) + \rho(A \cap B) \leq \rho(A) + \rho(B)$.*

For future convenience, we define a **loop** of M to be an element x of E such that $\rho(\{x\}) = 0$, and a **pair of parallel elements** of M to be a pair x, y of elements of E which are not loops, and for which $\rho(\{x, y\}) = 1$. A matroid which contains no loops or pairs of parallel elements is called a **simple matroid**.

The connection between matroid theory and graph theory may be seen by defining a matroid in terms of its circuits. We shall call a subset of E **dependent** if it is not independent, and a minimal dependent set will then be called a **circuit**. Since the circuits of a matroid determine the independent sets, we can redefine a matroid in terms of its circuits as follows:

A **matroid** $M = (E, \mathcal{C})$ consists of a non-empty finite set E , together with a collection \mathcal{C} of non-empty subsets of E (called **circuits**) satisfying the following properties:

- (C1) *no circuit properly contains another circuit;*
- (C2) *if C_1 and C_2 are distinct circuits each containing an element x , then there exists a circuit in $C_1 \cup C_2$ which does not contain x .*

(Note that this definition gives the extension to matroids of property (G1) (i).)

We conclude this section by defining a matroid on a set E in terms of a closure operation on the set $\mathcal{P}(E)$ of subsets of E . If $M = (E, \rho)$ is a matroid on E defined in terms of its rank function ρ , then the **closure** (or **span**) $c(A)$ of a subset A of E is defined to be the set of all those elements x of E which depend on A ; i.e., $c(A) = \{x \in E: \rho(A \cup \{x\}) = \rho(A)\}$. We can now redefine M in terms of c as follows:

A **matroid** $M = (E, c)$ consists of a non-empty finite set E , together with a function $c: \mathcal{P}(E) \rightarrow \mathcal{P}(E)$, satisfying the following properties.

- (ci) *for each subset A of E , $A \subseteq c(A) = c(c(A))$;*
- (cii) *if $A \subseteq B \subseteq E$, then $c(A) \subseteq c(B)$;*
- (ciii) *if $x \in c(A \cup \{y\})$, $x \notin c(A)$, then $y \in c(A \cup \{x\})$.*

The first two of these conditions express the fact that c is a **closure operation** on E , and the third says that c satisfies what is usually known as the **exchange condition**, which may be expressed informally by saying that if x depends on A and y , but not on A alone, then y depends on A and x . We leave it to the reader to verify that a subset A of E is independent in M if and only if no element of A depends on the others, i.e., if and only if $x \notin c(A - \{x\})$ for each $x \in A$.

A matroid M defined in terms of its closure operation is sometimes called a **pregeometry**; if in addition M is simple, then M is said to be a **geometry**. The reader who wishes to pursue matroid theory from this point of view should refer to Crapo and Rota's book [5] for a fuller treatment.

6. Some examples of matroids. To help the reader assimilate the various definitions given in the previous section, we now discuss several important types of matroid.

(1) **UNIFORM MATROIDS.** If E is a non-empty finite set, the **k -uniform matroid** on E is obtained by taking as bases all those subsets of E which contain exactly k elements. It follows immediately from this that the independent sets are precisely those subsets of E which contain not more than k elements, and that the rank of any subset A of E is either $|A|$ or k , whichever is the smaller. Of particular importance are the 0-uniform matroid on E (the **trivial matroid**) whose only independent set is the empty set and whose rank function is identically zero, and the $|E|$ -uniform matroid (the **discrete, or free, matroid**) in which every subset of E is independent and in which the rank of any subset of E is its cardinality. Note that the discrete matroid on E has only one base (namely E itself) and no circuits, and that the closure of any set A is A itself.

(2) **GRAPHIC MATROIDS.** As we indicated in Section 2, we can define a matroid on the set of edges of a graph G by taking as bases of the matroid the edges of the various spanning forests of G . We shall call this matroid the **circuit matroid** of G , and denote it by $M(G)$. It follows that a set of edges of G is independent if and only if it contains no circuit of G , and that the circuits of the matroid $M(G)$ are precisely the circuits of G . Note also that (by (G6)) the rank function of $M(G)$ is simply κ , and that if G is a simple graph, then $M(G)$ is a simple matroid.

We shall call a matroid M -**graphic** if M is isomorphic to the circuit matroid of some graph G . An example of a graphic matroid is obtained by letting $E = \{a, b, c\}$, and taking as independent sets \emptyset , $\{a\}$, $\{b\}$, $\{c\}$, $\{a, b\}$ and $\{a, c\}$ (see figure 6); an example of a nongraphic matroid is the 2-uniform matroid on a set of four elements, as the reader will discover if he tries to draw the graph. Graphic matroids have been characterized in terms of 'excluded minors' in an important series of papers by Tutte (see [31]); we shall present his characterization in Section 10.

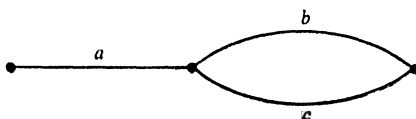


FIG. 6

(3) **COGRAPHIC MATROIDS.** The circuit matroid $M(G)$ is not the only interesting matroid which can be defined on the set E of edges of a graph G . In view of (G1) (ii), we can define a matroid in E by taking as its circuits the *cutsets* of G . This matroid is called the **cutset matroid** of G and is denoted by $M^*(G)$. Note that a set of edges of G is independent on $M^*(G)$ if and only if it contains no cutset of G .

We shall call a matroid M **cographic** if M is isomorphic to the cutset matroid of some graph G . The circuit matroids of the graphs K_5 and $K_{3,3}$ (depicted in figure 7) are examples of matroids which are not cographic; another example is given by the non-graphic matroid described above. We shall see later (in Sections 9,10) that the matroids $M(G)$ and $M^*(G)$ are dual matroids (in a sense to be made precise), and

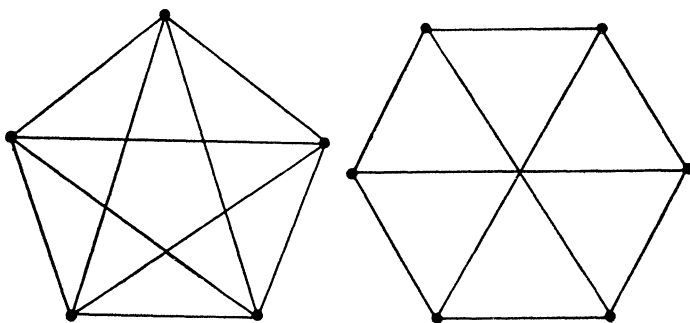


FIG. 7

that the concept of the planarity of a graph can be extended in a natural way to matroids.

(4) **REPRESENTABLE MATROIDS.** Let E be a finite set of vectors in some vector space V over a field F . We can define a matroid M on E by taking as independent sets of the matroid those subsets of E which are linearly independent in V . The bases of M are then precisely those subsets of E which span the same subspace as E (so that if the elements of E span V , then every base of M is a basis of V); note that the rank of any subset A of E is simply the dimension of the subspace of V spanned by A , and that the closure of A consists of all those elements of this subspace which lie in E .

We can similarly obtain a matroid from any finite set E of elements of a projective space over F ; in this case, the rank of any subset A of E is simply its rank $r(A)$, as defined in Section 3. The maximal subsets of rank one, two, three and $r(E) - 1$ are then called **points**, **lines**, **planes**, and **hyperplanes** respectively. (Note that if A is such a maximal set, then $c(A) = A$.)

We shall say that a matroid M on a set E is **linear over F** if M is isomorphic to a matroid obtained in the above way from a finite set of vectors in some vector space V over F . This definition, however, automatically excludes any matroid which contains too many loops or parallel elements, since each loop must correspond to the zero element of V , and parallel elements correspond to dependent vectors (leading to possible trouble if F is finite).

It is therefore convenient to define a matroid M to be **representable over a field F** (or, simply, **representable**) if there exists a (not necessarily one-one) *rank-preserving* mapping from E to the underlying set of some matroid which is linear over F . Note that this is equivalent to saying that the simple matroid M' obtained from M by removing all loops and identifying parallel elements is linear over F , or, in fact, that M' is isomorphic to a matroid defined in some projective space over F . It is also possible to define representability over a division ring, but we shall not discuss this here.

It turns out that some matroids are representable over every field (the so-called **regular matroids**), some over no fields (see Section 11), and some over only a restricted class of fields. Of special interest are the **binary matroids** which are representable over the field with two elements. It is not difficult to show that *every graphic matroid is binary*, and consequently we shall sometimes have to restrict ourselves to binary matroids when trying to extend properties of graphs to matroids.

(5) **ALGEBRAIC MATROIDS.** Let F be a field, and K be an extension of F . If E is a finite set of elements of K , a subset A of E is called dependent if the elements of A are algebraically dependent over F , i.e., if they satisfy a polynomial equation with coefficients in F . It can be proved that these dependent sets form the dependent sets of a matroid on E . Any matroid isomorphic to a matroid obtained in this way is called an **algebraic matroid**. At the time of writing, no one has yet proved the existence of matroids which are not algebraic.

(6) **TRANSVERSAL MATROIDS.** If E is a non-empty finite set and $\mathcal{S} = (S_1, \dots, S_m)$ is a family of non-empty subsets of E , then it follows from (T3), (T4) and (T6) that the partial transversals of \mathcal{S} may be taken as the independent sets of a matroid $M(\mathcal{S})$ on E . The bases of $M(\mathcal{S})$ are then the maximal partial transversals of \mathcal{S} , and the rank function is the function σ defined in Section 4.

We shall call a matroid M **transversal** if M can be obtained in the above way (for a suitable choice of E and \mathcal{S}). For example, the circuit matroid of the graph shown in figure 6 is a transversal matroid on $\{a, b, c\}$, since its independent sets are the partial transversals of the family $\mathcal{S} = (S_1, S_2)$, where $S_1 = \{a\}$, and $S_2 = \{b, c\}$; an example of a non-transversal matroid will be given later in this section. Note that if M is any k -uniform matroid on a set E , then M is a transversal matroid, since its independent sets are the partial transversals of the family $\mathcal{S} = (E, \dots, E)$ containing k E 's.

Transversal matroids have been characterized by Mason [16] and others, but there is as yet no known characterization in terms of excluded minors.

(7) **GAMMOIDS.** If E and Y are two disjoint sets of vertices in a directed graph* D , then we can define a matroid on E by taking as independent sets all those subsets A of E with the property that there exist $|A|$ directed paths, no two of which have any vertices in common, from A to a subset of Y (see figure 8). Any matroid obtained in this way is called a **gammoid**. Clearly gammoids may also be defined in terms of (undirected) graphs, although it is not known whether all gammoids can be obtained in this way. Note that a transversal matroid may be regarded as a gammoid in which the underlying directed graph is bipartite.

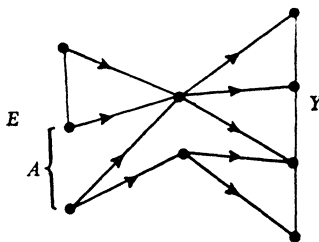


FIG. 8

(8) **THE FANO MATROID.** Of particular importance in matroid theory is the matroid F on the set $E = \{a, b, \dots, g\}$, in which the bases are all those subsets of E which contain exactly three elements, except $\{a, b, c\}$, $\{a, d, e\}$, $\{a, f, g\}$, $\{b, d, g\}$, $\{b, e, f\}$, $\{c, d, f\}$ and $\{c, e, g\}$. F is usually called the **Fano matroid**, and may be represented diagrammatically as in figure 9, where the bases are precisely those

* A **directed graph** (or **digraph**) is defined in the same way as a graph, except that the edges are now *ordered* pairs of vertices (v, w) . A **directed path** is then a finite sequence of distinct edges of the form $(v_0, v_1), (v_1, v_2), \dots, (v_{m-1}, v_m)$.

subsets of three elements which do not lie on a line. It can be proved that F is non-graphic, non-cographic and non-transversal; it can also be shown (see Section 11) that F is representable over any field of characteristic two, but not over any other fields.

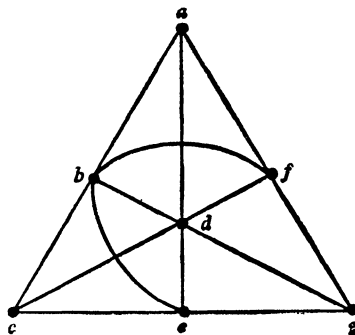


FIG. 9

(9) **UNIONS OF MATROIDS.** If $M_1 = (E_1, \rho_1)$ and $M_2 = (E_2, \rho_2)$ are two matroids, there are various ways of defining their union. For example, if E_1 and E_2 are disjoint, then their **disjoint union** is the matroid on $E_1 \cup E_2$ whose independent sets are obtained by taking the union of an independent set in M_1 and an independent set in M_2 ; note that the resulting rank function is then simply $\rho_1 + \rho_2$. As we shall see in Section 12, a similar construction may be used when $E_1 = E_2$, although in this case the rank function is rather more complicated.

(10) **FURTHER EXAMPLES.** In addition to the matroids already described, there are several other important types of matroid which we cannot discuss here. Among these are the matroids derived from simplicial complexes, and those associated with the partitions of a set. The reader will find both of these examples discussed in Crapo and Rota's book [5].

7. Matroids and lattices. There has recently been a lot of interest in the investigation of matroids from a lattice-theoretic point of view. In this section, we shall indicate how this connection between matroids and lattices arises; a fuller discussion will be found in Crapo and Rota [5].

If M is a matroid on a set E , and c is the closure operation on M described in Section 5, then we define a subset A of E to be a **closed set** (or **subspace**) if $c(A) = A$. For example, if M is a simple matroid, then the empty set and all subsets of E containing only one element are closed. It follows easily from properties (ci) and (cii) that if A and B are closed subsets of E , then so is $A \cap B$. It is not in general true, however, that their union $A \cup B$ is necessarily closed, but we can assert that $c(A \cup B)$ is always closed.

It is not difficult to prove from these remarks that if M is a matroid on a set E ,

then the closed subsets of E form a lattice $L(M)$ under set inclusion, in which the meet $A \wedge B$ of two closed sets A and B is equal to $A \cap B$, and their join $A \vee B$ is equal to $c(A \cup B)$. For example, the lattice corresponding to the 3-uniform matroid on the set $E = \{a, b, c, d\}$ is as shown in figure 10.

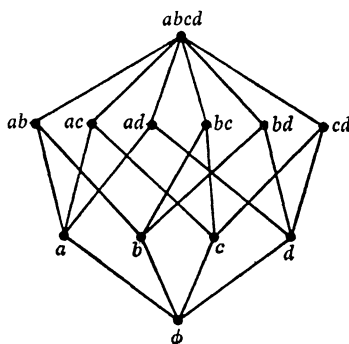


FIG. 10

It turns out that whatever matroid M we start with, the resulting lattice $L(M)$ is always a semimodular lattice. This means that the lattice has a height (i.e., rank) function r such that $r(A)$ measures the length of a maximal chain from the lowest element $c(\emptyset)$ to the element A of the lattice, and the function r satisfies the inequality $r(A \vee B) + r(A \wedge B) \leq r(A) + r(B)$, where A and B are any two elements of the lattice. Moreover, it can be shown that every element in the lattice can be expressed as the join of subsets of rank one. Such a lattice is usually called a **geometric lattice** (see Birkhoff [1]), and the elements of rank one, two, three and $r(E) - 1$ are called **points**, **lines**, **planes**, and **hyperplanes**, respectively. The reader should check that in the case of a linear matroid, these definitions agree with those given in the previous section. Note also that if M is a simple matroid, then the points of $L(M)$ are precisely the one-element subsets of E .

We have seen that if we are given a matroid M on a set E , then we can define a geometric lattice $L(M)$ whose elements are the closed subsets of E . It can also be shown that if L is any geometric lattice, then we can obtain a simple matroid $M(L)$ defined on the set E of points of L , by defining a subset A of E to be **independent** in $M(L)$ if and only if the join of the points in A has (lattice-) rank equal to $|A|$. It is not difficult to see that the resulting matroid $M(L)$ is the same as the simple matroid which gives rise to L in the manner described above. (The reader should verify this in the case of the lattice L of figure 10.)

It follows from this that there is a one-one correspondence between simple matroids and geometric lattices, and hence that matroid theory may be regarded essentially as the study of geometric lattices.

8. Duality in graph theory. In this section we shall describe various ways of

defining the dual of a graph. The idea of this is to motivate some of the results on matroid duality to be presented in the next section.

A graph G is called a **planar graph** if it can be embedded in the plane without crossings (in other words, the lines representing two edges of G are allowed to intersect only at a point of the plane which corresponds to a vertex to which they are both incident); any such embedding is then called a **plane graph**. (Strictly speaking, this should be called a 'plane embedding of a planar graph', but the abbreviation 'plane graph' is now standard.) The regions into which the edges of a plane graph divide the plane are called **faces**. For example, both of the graphs of figure 11 are planar, but only the second one is a plane graph; note that the second graph has four faces.

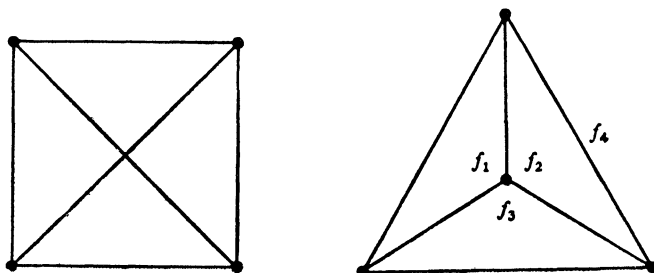


FIG. 11

If G is a plane graph, we form its **geometric-dual** G^* as follows: we take as the vertices of G^* one point inside each face of G , and then, for each edge e of G (adjoining faces f_1 and f_2 , say), we draw a corresponding edge e^* of G^* which crosses e (but no other edge of G) and joins those vertices of G^* which lie inside f_1 and f_2 . This procedure is illustrated in figure 12, with small crosses and dashed lines denoting the vertices and edges of G^* . It is a simple matter to check that if G is a connected plane graph, then its double-dual G^{**} is isomorphic to G .

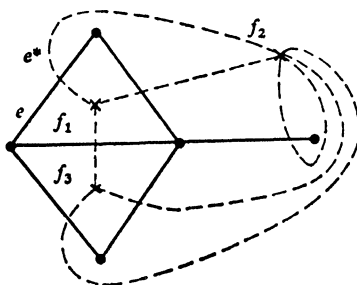


FIG. 12

If G is a planar graph, then a geometric-dual G^* of G may be defined by taking any plane embedding of G , and forming the geometric-dual in the manner described

above. It follows from this that a planar graph may have several different geometric-duals; all of them, however have certain properties in common—for example, the following two results always hold ([36] Section 15):

(G8) *A set of edges of G forms a circuit in G if and only if the corresponding set of edges of G^* forms a cutset in G^* .*

(G9) *If G is a planar graph, then G is bipartite if and only if G^* is Eulerian (i.e., the set of edges of G^* can be partitioned into disjoint circuits of G^*).*

It is obvious that a graph G is planar if and only if G has a geometric-dual, since we have not defined geometric-duality for non-planar graphs. What we should like to do is to find a definition of duality which generalizes the geometric-dual of a planar graph, but which also tells us whether or not a given graph is planar. One such definition, arising out of (G8), is as follows: a graph G^* is called an **abstract-dual** of a graph G if there is a one-one correspondence between the edges of G and those of G^* with the property that a set of edges of G forms a circuit of G if and only if the corresponding set of edges of G^* forms a cutset of G^* . The following results now hold (see [36] Section 15, and Parsons [24]):

(G10) *If G^* is an abstract-dual of G , then G is an abstract-dual of G^* .*

(G11) *A graph is planar if and only if it has an abstract-dual.*

An alternative definition of duality (which can be shown to be equivalent to the previous one) was given by Whitney. In this definition, \bar{H}^* denotes the graph obtained from G^* by removing the edges of H^* , and κ and γ are defined as in Section 2. We now define G^* to be a **Whitney-dual** of G if there is a one-one correspondence between the edges of G and those of G^* with the property that

$$\gamma(H) + \kappa(\bar{H}^*) = \kappa(G^*),$$

for any subgraph H of G whose vertex-set $V(H)$ is equal to the vertex-set of G . The following results can then be proved (see [36] Section 16):

(G12) *If G^* is a Whitney-dual of G , then G is a Whitney-dual of G^* .*

(G13) *A graph is planar if and only if it has a Whitney-dual.*

Although the definitions of abstract-duality and Whitney-duality may seem at first sight rather strange, they will turn out to be direct consequences of the definition of matroid duality.

9. Duality in matroids. Our aim in this section is to show how one can define matroid duality in such a way that the circuit and cutset matroids of a graph are duals of each other. The resulting definition will include the three definitions of duality

given in the previous section, and will also enable us to explain the similarity in graph theory between the properties of circuits and those of cutsets.

If M is a matroid on a set E , we define the **dual matroid** M^* to be the matroid on E whose bases are precisely the complements of the bases of M ; in other words, B^* is a base of M^* if and only if $E - B^*$ is a base of M . It is not difficult to check that this does in fact define a matroid, and that its rank function ρ^* is given by

$$\rho^*(A) = |A| + \rho(E - A) - \rho(E),$$

where $A \subseteq E$, and ρ is the rank function of M .

It follows immediately from this definition that (in contrast to the duality of planar graphs) every matroid has a dual, and that this dual is unique. Moreover, it is clear that the double-dual M^{**} is equal to M . It can also be proved without too much difficulty (see [36] Section 32) that if G is a graph, then the circuits of the dual matroid of $M(G)$ are precisely the cutsets of G , and hence that *the dual of $M(G)$ is simply $M^*(G)$* .

To see where this is all leading, let us define some 'co-notation'. If M^* is the dual of M , we define a **cocircuit** of M to be a circuit of M^* (so that, for example, the cocircuits of the circuit matroid $M(G)$ of a graph G are simply the cutsets of G). Similarly, we define a **cobase** of M to be a base of M^* , the **corank** of M to be the rank of M^* , and so on. Analogously, we say that a matroid is **cographic** if its dual is graphic, and in view of the remarks made above, this definition agrees with the one given in Section 6. (As an exercise, the reader may like to check that *the cocircuits of M are precisely the complements of the hyperplanes of M* .) The reason for introducing these extra definitions is that we need now deal only with the matroid M , instead of dealing with both M and M^* .

To illustrate this, let us first explain the similarity between the properties of circuits and the properties of cutsets in a graph (as illustrated by (G1), (G2) and (G7) of Section 2). The reason for this is simply that any result on the circuits of a matroid immediately gives us *two* results about graphs, since we can apply our matroid result either to the circuit matroid $M(G)$ of a graph G (giving us a result on the circuits of G), or to the cutset matroid $M^*(G)$ of G (giving us the corresponding result for the cutsets of G).

As an example of this, let us consider property (Gii) of Section 5. If we apply this to $M(G)$ we immediately deduce the result of (G1) (i); however, if we apply it to $M^*(G)$, then we obtain the result of (G1) (ii). In other words, the two results of (G1) are simply dual forms of a single result, and so, instead of having to prove two separate results in graph theory, we need prove only one result in matroid theory, and then use duality.

Another example is given by the results of (G7). It is easy to prove that *in any matroid, every cocircuit intersects every base* (since if C^* and B are a cocircuit and a base of M which are disjoint, then C^* is a circuit of M^* contained in a base $E - B$

of M^* , giving the required contradiction). On applying this result to the circuit matroid $M(G)$ of G , we immediately obtain (G7) (i); on applying it to the cutset matroid $M^*(G)$, we immediately obtain (G7) (ii). Note that both of these results may also be deduced from the dual of the above matroid result, namely that *in any matroid, every circuit intersects every cobase*.

It is probably worth pointing out at this stage that the result of (G2) does not generalize completely to arbitrary matroids. In fact, if C is any circuit and C^* is any cocircuit in a matroid M , then all we can say in general is that the number of elements common to C and C^* is not equal to one. However, if M is a binary matroid (as is the case with the circuit matroid of a graph), then it turns out that the number of elements common to C and C^* is even, generalizing (G2). Since the converse of this is also true, it follows that *the dual of a binary matroid is binary*.

We have just seen how matroid duality can be used to give us greater insight into problems involving the circuits and cutsets of a graph. We now show briefly how the results of Section 8 fit in with the definition of matroid duality.

We note first that the matroid definition of duality generalizes the definition of an abstract-dual of a graph G ; this is clear since the circuits of $M(G)$ correspond to the cocircuits of $M^*(G)$ and hence the circuits of G correspond to the cutsets of G^* . Since the abstract-dual of a planar graph is equivalent to the Whitney-dual and generalizes the geometric-dual, it follows that these other two definitions of duality are also consequences of the matroid definition of duality. Note that although a planar graph may have several different geometric-duals, they all give rise to the same matroid. Note also that the rather artificial-looking equation in the definition of the Whitney-dual is simply a restatement of the expression for ρ^* given at the beginning of this section, and that the equation $M^{**} = M$ is the natural matroid generalization of properties (G10) and (G12).

We conclude this section by remarking that (G9) has a natural extension to binary matroids which takes the following form: *if M is a binary matroid, then M is bipartite if and only if M^* is Eulerian*; in this statement, a matroid is called **bipartite** if every circuit contains an even number of elements, and is called **Eulerian** if E can be expressed as the union of disjoint circuits. The reader is referred to Welsh [33] for a proof of this result.

10. Graphic Matroids. In this section we shall present Tutte's fundamental result on the characterization of graphic matroids [31]. We shall find it convenient, however, to start with a few remarks about planar graphs.

If G is any graph, then we can obtain new graphs from G by any succession of the following two operations:

- (i) deleting one or more of its edges;
- (ii) contracting one or more of its edges, i.e., removing an edge $e = \{v, w\}$ and identifying the vertices v and w in such a way that all edges which were formerly

incident to either v or w are now incident to the new vertex (see figure 13).

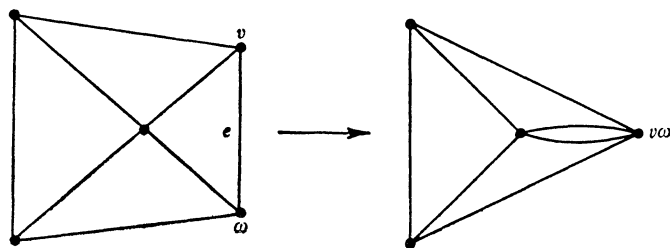


FIG. 13

In 1930, Kuratowski [14] obtained a characterization of planar graphs which has been expressed by Harary and Tutte [9] in the following form (see [36] Section 12):

KURATOWSKI'S THEOREM. *A connected graph is planar if and only if it cannot be reduced to either K_5 or $K_{3,3}$ (see figure 7) by any succession of the above operations.*

The operations of deletion and contraction of edges in a graph have analogues in matroid theory. If M is a matroid on a set E , and A is a subset of E , then the **deletion matroid** (or **restriction matroid**) $M \times A$ is the matroid on A whose circuits are precisely those circuits of M which are contained in A ; similarly, the **contraction matroid** $M \cdot A$ is the matroid on A whose cocircuits are precisely those cocircuits of M which are contained in A . We leave it to the reader to verify that if M is the circuit matroid of a graph G , then these matroids correspond to the graphs obtained by the operations described above. Any matroid which is obtained from M by a succession of deletions and contractions is called a **minor** of M .

It is a simple matter to show that if a matroid M is binary, graphic or cographic, then any minor of M has the same property. It follows from Section 6 that if M is a graphic matroid, then M is binary and contains no minor which is isomorphic to $M^*(K_5)$, $M^*(K_{3,3})$, F or F^* , where F denotes the Fano matroid.

In 1958, Tutte proved in an important series of papers that these conditions were not only necessary conditions for a matroid to be graphic, but were also sufficient, thus characterizing graphic matroids. He also proved the very deep result that a *binary matroid M is regular (see Section 6(4)) if and only if it contains no minor isomorphic to F or F^** . We can thus state Tutte's result in the following form:

TUTTE'S THEOREM. *A matroid M is graphic if and only if it is regular and contains no minor isomorphic to $M^*(K_5)$ or $M^*(K_{3,3})$.*

On applying Tutte's theorem to M^* , and using the fact that the dual of a regular matroid is regular, we immediately obtain a characterization of cographic matroids:

THEOREM. *A matroid M is cographic if and only if it is regular and contains no minor isomorphic to $M(K_5)$ or $M(K_{3,3})$.*

Since a graph is planar if and only if it has a dual, it seems reasonable to define a matroid M to be **planar** if both M and its dual M^* correspond to graphs, in other words if M is both graphic and cographic. In view of the above remarks, we can now give a characterization of planar matroids which corresponds to Kuratowski's characterization of planar graphs:

THEOREM. *A matroid is planar if and only if it is regular and contains no minor isomorphic to $M(K_5)$, $M(K_{3,3})$ or their duals.*

11. Representable matroids. In this section we discuss briefly the representability of matroids from a more geometrical point of view. In what follows, every matroid M will be assumed simple; it follows from this that if M is representable, then M must be isomorphic to a matroid defined in some projective space over some field, and hence that the configurations which arise must not contradict well-known results in projective geometry.

To illustrate this, we first consider Pappus' theorem which states that if in figure 14, the points $\{a, b, c\}$ and the points $\{d, e, f\}$ are collinear, then so are the points $\{g, h, i\}$. It follows that if we take M to be the matroid on $\{a, b, \dots, i\}$ whose bases are all those subsets containing three elements *except*

$\{a, b, c\}$, $\{a, e, g\}$, $\{a, f, h\}$, $\{b, d, g\}$, $\{b, f, i\}$, $\{c, d, h\}$, $\{c, e, i\}$ and $\{d, e, f\}$,

then M cannot be representable over any field, since if it were, then by Pappus' theorem the set $\{g, h, i\}$ would also have to be a dependent set, contradicting the fact that $\{g, h, i\}$ is a base of M .

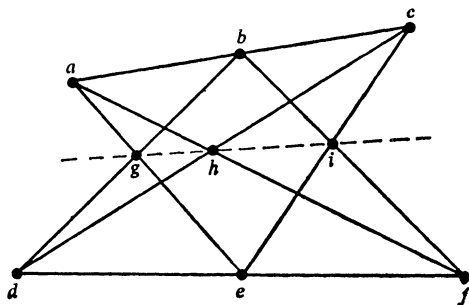


FIG. 14

A similar situation holds with Desargues' theorem, which states that if in figure 15 the triangles bcd and efg are in perspective from the point a (i.e., the points $\{a, b, e\}$ are collinear, as are the points $\{a, c, f\}$ and $\{a, d, g\}$), then the points $\{h, i, j\}$ are also collinear. It follows as before that if M is the matroid on $\{a, \dots, j\}$ whose

bases are all those subsets containing three elements *except*

$$\{a, b, e\}, \{a, c, f\}, \{a, d, g\}, \{b, c, h\}, \{b, d, i\}, \{c, d, j\}, \{e, f, h\}, \{e, g, i\}$$

and $\{f, g, j\}$,

then M cannot be representable over any field.

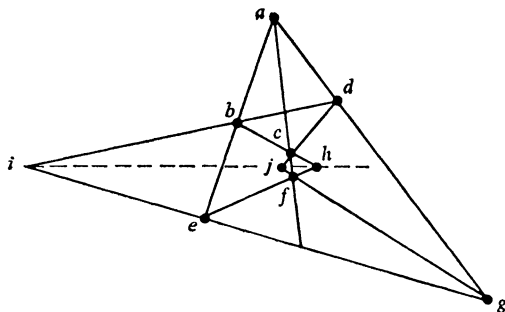


FIG. 15

Before leaving the subject of representability, we return briefly to the Fano matroid F which corresponds (see figure 9) to the seven-point projective plane. To investigate the representability of F , we assign plane projective coordinates to each point, choosing (as we may)

$$a = (1, 0, 0), \quad c = (0, 1, 0), \quad g = (0, 0, 1) \quad \text{and} \quad d = (1, 1, 1);$$

it follows immediately that $b = (1, 1, 0)$, $e = (0, 1, 1)$ and $f = (1, 0, 1)$. But these three points are collinear if and only if the determinant formed by their coordinates is zero, i.e., if and only if $2 = 0$. It follows that the Fano matroid is representable only over fields of characteristic two. It also follows from this discussion that if F' is the matroid which corresponds to figure 9 with the line joining the points $\{b, e, f\}$ removed, then F' is representable over all fields *except* those of characteristic two. Note that the disjoint union of F and F' is not representable over any field.

For further results of this kind, and a more complete discussion of representability in general, the reader is referred to MacLane [15], or to the survey article by Ingleton [12].

12. Transversal matroids. Up to now, we have been concerned primarily with the relationship between matroid theory and graph theory. We now indicate briefly how matroid theory can be used to prove results in transversal theory. This subject is dealt with in far greater depth in Mirsky's book [18], and the reader is referred there for further details.

We recall that if $\mathcal{S} = (S_1, \dots, S_m)$ is a family of non-empty subsets of a finite set

E , then the partial transversals of \mathcal{S} form the independent sets of a matroid $M(\mathcal{S})$ on E . Using this fact we can give a very elementary matroid proof of property (T2) (of Section 4). It is clear that if \mathcal{S} has a transversal containing A , then the properties (i) and (ii) of (T2) hold; conversely if (i) and (ii) hold, then by (ii), A is an independent set in $M(\mathcal{S})$ and hence can be extended to a base of $M(\mathcal{S})$. The result now follows from (i), since every base of $M(\mathcal{S})$ is a transversal of \mathcal{S} , proving the result.

There are several other proofs in transversal theory which can be simplified using matroids, and one may ask whether there is a generalization of Hall's theorem (see Section 4) which gives necessary and sufficient conditions for the existence of an independent transversal of a family \mathcal{S} of subsets of a finite set E with a matroid M defined on it. Such a generalization was given by R. Rado [27] in 1942:

RADO'S THEOREM. *Let E be a finite set, M a matroid on E , and $\mathcal{S} = (S_1, \dots, S_m)$ be a family of non-empty subsets of E . Then \mathcal{S} has a transversal which is independent in M if and only if, for each k satisfying $1 \leq k \leq |E|$, the union of any k of the subsets S_i , has rank at least k .*

Note that if M is the discrete matroid on E , then Rado's theorem reduces to Hall's theorem.

Rado's theorem has a wide variety of applications in transversal theory. One of the most well-known of these (see [36] Sections 27, 33) is the problem of finding necessary and sufficient conditions for two families \mathcal{S} and \mathcal{T} of subsets of a given set E to share a common transversal. This problem is of some importance in the study of timetabling as may be seen by taking, for example, the underlying set E to be the set of hours when mathematics lectures can be given, the family \mathcal{S} to consist of the sets of hours when each of the professors is able to lecture, and the family \mathcal{T} to consist of the hours when each of the classrooms is available. A common transversal then gives us a way of assigning professors to rooms at times when both are available.

Rado's theorem has also been used (by Welsh [34]) to obtain results on the union of matroids. If M_1, \dots, M_k are matroids on the same set E , with rank functions ρ_1, \dots, ρ_k respectively, then we can define a new matroid $M_1 \cup \dots \cup M_k$ on E by taking as independent sets all possible unions of the form $A_1 \cup \dots \cup A_k$, where A_i is independent in M_i for each i . That this actually defines a matroid seems to have been first pointed out by Nash-Williams [23], and its rank function $\tilde{\rho}$ is given by

$$\tilde{\rho}(A) = \min_{X \subseteq A} \{ \rho_1(X) + \dots + \rho_k(X) + |A - X| \}, \quad (A \subseteq E).$$

In particular, if $M_1 = \dots = M_k (= M, \text{ say})$, then the rank function simplifies to

$$\tilde{\rho}(A) = \min_{X \subseteq A} \{ k \rho(X) + |A - X| \},$$

where ρ is the rank function of M .

This last result has been used by Edmonds [7], and others, to give simple proofs

of several deep results in graph theory, transversal theory, and vector spaces. For example, if M is the circuit matroid of a graph G , then G contains k edge-disjoint spanning forests if and only if M contains k disjoint bases, i.e., if and only if the union of k copies of M has rank at least $k\rho(E)$. This can be restated as follows:

THEOREM. *A graph G contains k edge-disjoint spanning forests if and only if, for each subgraph H of G , $k(\kappa(G) - \kappa(H)) \leq m(G) - M(H)$, where $m(H)$ denotes the number of edges in H .*

In a similar way, the above result can be used to give a necessary and sufficient condition for G to split into at most k forests. In this case, the rank of the union of k copies of M is simply $|E|$, and the result takes the following form:

THEOREM. *A graph G can be split up into at most k forests if and only if, for each subgraph H of G , $k \cdot \kappa(H) \geq m(H)$.*

Although they are easy deductions from our results on the union of matroids, both of these theorems are very difficult to prove by straightforward graph-theoretic techniques (see the papers of Tutte [30] and Nash-Williams [21], [22]).

We conclude this section by stating a theorem of Horn [11] on vector spaces, which can be deduced in exactly the same way as the second of the above theorems.

THEOREM. *If E is a finite set of vectors in a vector space, then E can be divided into k disjoint linearly independent subsets if and only if, for each subset A of E , $|A| \leq k \cdot \text{rank } A$, where the rank of A is as defined in Section 3.*

13. Some recent results. In this final section, we shall discuss a few results in matroid theory which have been proved in the last two or three years.

(1) **THE ENUMERATION OF MATROIDS.** It is easy to see that the number $f(n)$ of non-isomorphic matroids on a set E of n elements cannot exceed 2^{2^n} , since E has 2^n subsets, and each of these may be dependent or independent. What is surprising is that this seemingly rather crude upper bound turns out to be very sharp, since Piff and Welsh [26] have proved that for any value of $\lambda < 1$, $f(n)$ is bounded below by $2^{2^{\lambda n}}$, if n is sufficiently large. In fact, they managed to obtain the following sharper bounds for n sufficiently large:

$$2^\alpha < f(n) < 2^\beta,$$

where $\alpha = 2^n/n^{\frac{1}{2}}$ and $\beta = 2^n/n^{\frac{1}{4}}$.

If we now restrict ourselves to transversal matroids, and let $t(n)$ denote the number of non-isomorphic transversal matroids on a set E of n elements, then each of the subsets of E in the family \mathcal{S} can be chosen in 2^n ways, and hence $t(n)$ is bounded above by $(2^n)^n = 2^{n^2}$. Stricter bounds have been obtained by Heron and Piff

(unpublished) who have proved that

$$2^{\frac{1}{2}n^2} \leq t(n) \leq 2^{\frac{3}{2}n^2}.$$

Little is known about the number of non-isomorphic graphic or representable matroids.

(2) CONNECTIONS BETWEEN VARIOUS TYPES OF MATROIDS. It has recently been shown by Piff and Welsh [25] that if M_1 and M_2 are matroids on the same set, both of which are representable over a field F , then their union $M_1 \cup M_2$ is also representable over F , provided that F has sufficiently large cardinality. It follows, using the easily-proved fact that any transversal matroid can be expressed as the union of matroids of rank one, that *every transversal matroid is representable over all sufficiently large fields* (and, in particular, over every infinite field). Since every k -uniform matroid is transversal, it follows from this that every k -uniform matroid is representable over all sufficiently large fields.

It has also been proved recently by de Sousa and Welsh [6] that *a transversal matroid is binary if and only if it is graphic*. Related to this is the result of Bondy [2] that *the circuit matroid $M(G)$ of a graph G is transversal if and only if G contains no subgraph homeomorphic to K_4 or C_n^2 for some n* . (In this statement, K_4 denotes either of the graphs shown in figure 11, and C_n^2 is the graph obtained by 'doubling up' the edges of an n -gon (see Figure 16); the theorem then states that no subgraph of G can be obtained by inserting new vertices into the edges of either K_4 or C_n^2 .)

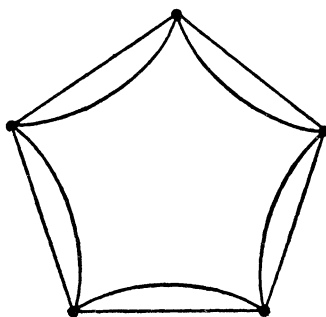


FIG. 16

(3) PRESENTATIONS OF TRANSVERSAL MATROIDS. If M is a transversal matroid on a set E , whose independent sets are the partial transversals of a family \mathcal{S} of subsets of E , then \mathcal{S} is called a **presentation** of M . It is not difficult to show that *if M has rank r , then there exists a presentation of M which contains only r subsets*. Moreover, it has been shown by Bondy and Welsh [3] that these subsets can all be taken to be cocircuits of M .

(4) RESULTS ON GAMMOIDS. It is not difficult to show that the dual of a transversal matroid is not necessarily transversal, and it is therefore a worthwhile question to ask what the duals of transversal matroids look like. Since every transversal matroid is a gammoid, one can also ask what the duals of gammoids look like. This problem has been solved recently by Mason [17] and Ingleton and Piff [13], who have shown that the dual of a gammoid is always a gammoid, and hence that the dual of a transversal matroid is also a gammoid. Moreover, Ingleton and Piff have shown that there are some important gammoids (called 'strict gammoids') which have certain natural properties, and which turn out to be precisely the duals of transversal matroids.

Ingleton and Piff also showed that *a matroid M is a gammoid if and only if M is a contraction of a transversal matroid*. By imitating the argument of Piff and Welsh, they were able to prove that *every gammoid is representable over all sufficiently large fields*.

(5) THE CRITICAL PROBLEM. If M is a matroid on a set E , then the critical problem as posed by Crapo and Rota [5] is the problem of determining the minimum number k of hyperplanes H_i of M , such that $H_1 \cap \dots \cap H_k = \emptyset$.

By duality, this problem is equivalent to the problem of finding the minimum number k of circuits of M whose union is E , although the problem loses much of its geometrical significance when expressed in this way. Recently the critical problem has been solved for quite a number of matroids, although there is some way to go before any results of great significance are obtained. The importance of the critical problem stems mainly from the fact that several of the famous unsolved problems in the coloring of graphs (including the celebrated four-color conjecture) may be shown to be special cases of the critical problem. Crapo and Rota have expressed the hope that by developing techniques for solving the critical problem in a few simple cases, one may eventually find a suitable approach for tackling these coloring problems.

References

1. G. Birkhoff, *Lattice Theory*, 3rd ed. Amer. Math. Soc. Colloq. Publ., 25 (1967).
2. J. A. Bondy, 'Transversal matroids, base-orderable matroids, and graphs, *Quart. J. Math. Oxford* 23 (1972) 81-89.
3. J. A. Bondy and D. J. A. Welsh, Some results on transversal matroids and constructions for identically self-dual matroids, *Quart. J. Math. Oxford Ser.*, 22 (1971) 435-451.
4. R. A. Brualdi and E. B. Scrimger, Exchange systems, matchings, and transversals. *J. Combinatorial Theory*, 5(1968) 244-257.
5. H. H. Crapo, and G. -C. Rota, *Combinatorial Geometries*, M.I.T. Press, 1971.
6. J. de Sousa and D. J. A. Welsh, A characterisation of binary transversal structures, (to appear).
7. J. Edmonds, Minimum partition of a matroid into independent subsets, *J. Res. Nat. Bur. Standards*, 69 B, (1965) 67-72.
8. J. Edmonds and D. R. Fulkerson, Transversals and matroid partition, *J. Res. Nat. Bur. Standards* 69 B, (1965) 147-153.

9. F. Harary and W. T. Tutte, A dual form of Kuratowski's theorem, *Canad. Math. Bull.*, 8 (1965) 17–20, 373.
10. F. Harary and D. J. A. Welsh, Matroids versus graphs, in *The Many Facets of Graph Theory*, Springer Lecture Notes 110, 1969.
11. A. Horn, A characterization of unions of linearly independent sets, *J. London Math. Soc.*, 30 (1955) 494–496.
12. A. W. Ingleton, Representation of matroids, in *Combinatorial Mathematics and its Applications*, Academic Press, New York, 1971.
13. A. W. Ingleton and M. J. Piff, Gammoids and transversal matroids, (to appear).
14. K. Kuratowski, Sur le problème des courbes gauches en topologie, *Fund. Math.*, 15 (1930) 271–283.
15. S. MacLane, Some interpretations of abstract linear dependence in terms of projective geometry, *Amer. J. Math.*, 58 (1936) 236–240.
16. J. H. Mason, A characterization of transversal independence spaces, in *Théorie des Matroides*, Springer Lecture Notes 211, 1971.
17. ———, On a class of matroids arising from paths in graphs, *Proc. London Math. Soc.*, 25 (1972) 55–74.
18. L. Mirsky, *Transversal Theory*, Academic Press, New York, 1970.
19. ———, Transversal theory and the study of abstract independence, *J. Math. Anal. Appl.*, 25 (1969) 209–217.
20. L. Mirsky and H. Perfect, Applications of the notion of independence to problems of combinatorial analysis. *J. Combinatorial Theory*, 2 (1967) 327–357.
21. C. St. J. A. Nash-Williams, Edge-disjoint spanning trees of finite graphs, *J. London Math. Soc.*, 36 (1961) 445–450.
22. ———, Decomposition of finite graphs into forests, *J. London Math. Soc.*, 39 (1964) 12.
23. ———, An application of matroids to graph theory, *Proc. Symp. Rome*, Dunod (1966) 263–265.
24. T. D. Parsons, On planar graphs, this MONTHLY, 78 (1971) 176–178.
25. M. J. Piff and D. J. A. Welsh, On the vector representation of matroids, *J. London Math. Soc.*, 2 (1970) 284–288.
26. ——— and ———, On the number of combinatorial geometries, *Bull. London Math. Soc.*, 3 (1971) 55–56.
27. R. Rado, A theorem on independence relations, *Quart. J. Math. (Oxford)* 13 (1942) 8 3–89.
28. ———, Axiomatic treatment of rank in infinite sets, *Canad. J. Math.*, 1 (1949) 337–343.
29. W. T. Tutte, *Introduction to the Theory of Matroids*, American Elsevier, New York, 1971.
30. ———, On the problem of decomposing a graph into n connected factors, *J. London Math. Soc.*, 36 (1961) 221–230.
31. ———, Lectures on matroids, *J. Res. Nat. Bur. Stand.*, 69B (1965) 1–47.
32. B. L. van der Waerden, *Moderne Algebra*, 2nd. ed., Springer, Berlin, 1937.
33. D. J. A. Welsh, Euler and bipartite matroids, *J. Combinatorial Theory*, 6 (1969) 375–377.
34. ———, On matroid theorems of Edmonds and Rado, *J. London Math. Soc.*, 2 (1970) 251–256.
35. H. Whitney, On the abstract properties of linear dependence, *Amer. J. Math.*, 57 (1935) 509–533.
36. R. J. Wilson, *Introduction to Graph Theory*, Oliver & Boyd, (Edinburgh) and Academic Press (New York) 1972.

STOCHASTIC EQUATIONS AND THEIR APPLICATIONS

G. C. PAPANICOLAOU, Courant Institute, New York University

1. Introduction. Almost all problems in physics, engineering, economics, biology, and other sciences to which mathematical methods are applicable are basically stochastic rather than deterministic. It is more appropriate to attempt to determine the probabilities with which the phenomena under investigation occur rather than the precise phenomenon which occurs. Nevertheless, the majority of mathematical methods are based on deterministic models. This is a reasonable first approximation which frequently renders the problem mathematically tractable. It is also quite adequate for many problems. However, the errors due to stochastic effects may accumulate in prolonged observations of a phenomenon so that the deterministic analysis becomes useless. It is necessary therefore to have a rational method for taking into consideration stochastic effects.

The problem of accumulation of error in random phenomena has been central in the development of probability theory. Until recently, however, investigations have been limited to simple situations which exclude several problems of interest. Let us consider some examples.

Suppose we wish to measure a physical quantity such as temperature at a fixed time and location. The measuring process is subject to errors which are due to many individually negligible causes. Thus by invoking the classical Central Limit Theorem [4] we conclude that the error in measurement is a Gaussian random variable with zero mean and variance chosen to fit available data or estimated in some other way. This is the well-known theory of errors in its simplest form.

Let us also consider the following problem. Let $u(t)$ represent a physical quantity as a function of time and suppose that $u(t)$ satisfies, for example, the differential equation

$$\ddot{u}(t) + a(t)\dot{u}(t) + b(t)u(t) = 0, \quad u(0) = u_0,$$

$$\dot{u}(0) = \dot{u}_0, \text{ where } f' \equiv \frac{d}{dt}f.$$

The functions $a(t)$ and $b(t)$ represent properties of the dynamical system determining the evolution of $u(t)$. Let us assume that $a(t)$ and $b(t)$ are random functions. Then $u(t)$ is a random function defined by a random or stochastic ordinary differential equation. If $a(t)$ and $b(t)$ fluctuate little from their expected values which we denote by $E\{a(t)\}$, $E\{b(t)\}$, then, as a first approximation we may solve the deterministic problem:

George Papanicolaou received his Ph. D. at the Courant Institute in 1969 under J. B. Keller. Since then he has held positions at the Univ. Heights Campus of N. Y. U. and at the Courant Institute. His main research is in the subject of this article. *Editor.*

$$\ddot{w}(t) + E\{a(t)\}\dot{w}(t) + E\{b(t)\}w(t) = 0, \quad w(0) = u_0, \quad \dot{w}(0) = \dot{u}_0.$$

The question arises: how good is this approximation? In particular, how does $w(t)$ compare with $E\{u(t)\}$? For short times we do not expect much discrepancy. This is the justification for considering deterministic problems. But at large times, $E\{u(t)\}$ may deviate very significantly from $w(t)$. A basic problem in the study of stochastic equations, and the one we shall consider here, is to find effective methods for computing the statistical characteristics of the solution given the statistical characteristics of the coefficients and of the initial or boundary conditions.

Let us observe that the simple theory of errors described above is not adequate for the study of stochastic equations. The methods of modern probability theory can be used, however, to obtain a number of interesting results on the behavior of solutions of stochastic equations. We shall present some of these results here.

In section 2, we give a few examples of physical problems that lead to stochastic ordinary differential equations. We limit ourselves mainly to stochastic ordinary differential equations because they are the simplest examples. Relatively little progress has been achieved for more general problems, such as stochastic partial differential equations, despite the fact that many physical problems lead naturally to such equations.

In section 3, we formulate the basic problems associated with stochastic ordinary differential equations, in the language of the theory of probability.

In section 4 we deal with a special class of problems: equations whose coefficients are Markov processes. We include here some applications, which are of independent interest, to singular perturbation of deterministic partial differential equations.

The basic limit theorem of Stratonovich [25] and Has'minskii [9] is presented in section 5. This is a far reaching generalization of the classical Central Limit Theorem and gives satisfactory results for many problems. This result was introduced in [26] and was applied extensively there. It was also obtained in a different manner in [14, 15, 21]. An operator theoretic proof is given in [22].

In section 6 we apply the theorem of section 5 to two examples.

2. Examples of problems leading to stochastic differential equations. Our first example is the random harmonic oscillator. Let $u(t)$ denote the displacement of a particle of mass m from its equilibrium position and let a linear spring with spring constant k connect the particle with a fixed support. Then $u(t)$ satisfies the equation of motion

$$(2.1) \quad m\ddot{u}(t) + ku(t) = f(t), \quad u(0) = u_0, \quad \dot{u}(0) = \dot{u}_0.$$

Here $f(t)$ is an external force acting on the particle and u_0, \dot{u}_0 are the initial displacement and velocity. If $f(t)$ is a random function of time then $u(t)$ will be a random function of time. (We delay a precise mathematical statement of this problem and the others considered in this section until section 3.) Similarly, u_0 and \dot{u}_0 may be

random variables. Under any circumstances the determination of the statistical characteristic of $u(t)$ is relatively simple since (2.1) is solved explicitly by

$$(2.2) \quad u(t) = u_0 \cos \sqrt{\frac{k}{m}} t + \sqrt{\frac{m}{k}} \dot{u}_0 \sin \sqrt{\frac{k}{m}} t + \sqrt{\frac{m}{k}} \int_0^t \frac{f(s)}{m} \sin \sqrt{\frac{k}{m}} (t-s) ds.$$

Thus, for example, $E\{u(t)\}$, the expected value of $u(t)$, can be determined by taking expected values on both sides of (2.2). Other statistical characteristics can be obtained in the same direct manner but computations may become lengthy.

Let us consider briefly the mean or expected energy associated with the oscillator. Let us assume that

$$(2.3) \quad E\{u_0\} = E\{\dot{u}_0\} = E\{f(t)\} = 0,$$

that u_0 , \dot{u}_0 , $f(t)$ are independent, and that $f(t)$ is a stationary random process with covariance

$$(2.4) \quad E\{f(t)f(s)\} = \sigma^2 R(t-s).$$

We also assume that

$$(2.5) \quad l = \int_0^\infty R(\sigma) d\sigma < \infty.$$

The quantity l has the dimension of time and is called the **correlation time** of $f(t)$. Now upon multiplying (2.1) by $\dot{u}(t)$, integrating from 0 to t , and taking expected values, we obtain

$$(2.6) \quad E \left\{ \frac{m}{2} \dot{u}^2(t) + \frac{k}{2} u^2(t) \right\} - E \left\{ \frac{m}{2} \dot{u}_0^2 + \frac{k}{2} u_0^2 \right\} = \int_0^t E\{\dot{u}(s)f(s)\} ds.$$

The quantity $\frac{1}{2} m \dot{u}^2 + \frac{1}{2} k u^2$ is the instantaneous energy of the oscillator. To compute the change in time of the mean energy from (2.6) we compute the right-hand side using our hypotheses (2.3)–(2.5) and (2.2):

$$(2.7) \quad \begin{aligned} \int_0^t E\{\dot{u}(s)f(s)\} ds &= \frac{\sigma^2}{m} \int_0^t \int_0^s R(s-\sigma) \cos \sqrt{\frac{k}{m}} (s-\sigma) d\sigma ds \\ &= \frac{\sigma^2}{m} \int_0^t (t-\sigma) R(\sigma) \cos \sqrt{\frac{k}{m}} \sigma d\sigma. \end{aligned}$$

Thus from (2.5), (2.7) and (2.6) we conclude that when t is large, the mean energy of the oscillator increases linearly with t .

This conclusion should be contrasted with the deterministic case when $f(t)$ is periodic. Here the energy is bounded if the period of f is different from $2\pi\sqrt{k/m}$ and increases like t^2 when f is periodic with this period (resonance).

Problem (2.1) is uncharacteristically simple because its solution is known explicitly. A more interesting problem arises when either the "constants" of the oscillator m and k are random functions, or when the spring is nonlinear, or when both situations arise. Let us suppose that the spring "constant" $k(t)$ is a random function so that we have

$$(2.8) \quad m\ddot{u}(t) + k(t)u(t) = 0, \quad u(0) = u_0, \quad \dot{u}(0) = \dot{u}_0.$$

If we replace $k(t)$ by its expected value in (2.8), this amounts to taking expected values and then assuming that

$$(2.9) \quad E\{k(t)u(t)\} = E\{k(t)\}E\{u(t)\}.$$

But we have no way of knowing whether (2.9) is valid or not since $u(t)$ is itself a functional of $k(\sigma)$ for $\sigma \in [0, t]$. In fact (2.9) is false in general.

Let us suppose that $k(t)$ is a stationary process so that

$$(2.10) \quad E\{k(t)\} = k_0 > 0, \quad E\{(k(t) - k_0)(k(s) - k_0)\} = \sigma^2 R(t - s),$$

$$l = \int_0^\infty R(\sigma) d\sigma < \infty.$$

The correlation time l is representative of the time length over which the values of $k(t)$ are significantly correlated. There is another natural time scale associated with (2.8) namely the period

$$(2.11) \quad \nu = 2\pi \sqrt{\frac{m}{k_0}}$$

of oscillation when $k(t)$ is replaced by k_0 . We now distinguish the following three cases for special consideration, when the variance of the inhomogeneities σ^2 is small.

(i) $\nu \ll l$. In this case replacing $k(t)$ by its mean value k_0 will not have appreciable effect over many periods of oscillation. The motion will be approximated well by the deterministic equation.

(ii) $\nu \gg l$. In this case we expect that stochastic effects are quite important at times of the order of magnitude of several periods. In section 5, we shall see that in this case a generalized Central Limit Theorem is valid.

(iii) ν comparable to l . In this case the precise nature of the process $k(t)$ will be important. If we assume that $k(t)$ is Markovian then sometimes it may be possible to obtain information about $u(t)$. Equations with Markovian coefficients are discussed in section 4.

Let us now suppose that the oscillator is nonlinear and is acted upon by a random force so that

$$(2.12) \quad m\ddot{u} + V'(u(t)) = f(t), \quad u(0) = u_0, \quad \dot{u}(0) = \dot{u}_0, \quad V' = \frac{dV}{du}.$$

Here $V(u)$ is the potential energy function of the spring and is assumed to be convex upwards in the neighborhood of $u = 0$. Let us also assume that $f(t)$ is stationary with mean zero and (2.4), (2.5) hold. We may again consider the three regimes encountered above. Now we may take for v the quantity $2\pi\sqrt{m/V''(0)}$ provided V is a smooth function near $u = 0$.

Another example of interest is the following. Let $u(t, x)$ denote the transverse displacement of a taut string occupying the interval $[0, L]$. Then u satisfies the equation:

$$(2.13) \quad \begin{aligned} \rho u_{tt} &= F u_{xx}, \quad u(t, 0) = \cos \omega t, \quad u(t, L) = 0, \\ u(0, x) &= f(x), \quad u_t(0, x) = g(x). \end{aligned}$$

We have assumed that the left end point of the string is being subjected to oscillations of frequency ω and we have denoted by $\rho = \rho(x) > 0$ the mass density of the string and by F the tensile force. Let us suppose that $\rho(x)$ is a random function of position with mean $\rho_0 > 0$ and sufficiently small fluctuation. If we ignore the transient behavior of the string and examine only the steady state solution

$$u(t, x) = \text{Real part of } [e^{i\omega t} v(x)],$$

then $v(x)$ satisfies the boundary value problem

$$(2.14) \quad v_{xx} + \frac{\omega^2}{F} \rho(x) v = 0, \quad v(0) = 1, \quad v(L) = 0.$$

Thus $v(x)$ is a random function of position. Problem (2.14) is more difficult than (2.8) because it is a boundary value problem. We may again consider the three cases we introduced following (2.11). But for case (ii) and case (iii) no results analogous to the ones for initial value problems are known.

Stochastic eigenvalue problems also arise frequently. For example, let $u(t, x)$ be again as above but now suppose both end points are kept fixed. Then we have

$$(2.15) \quad \begin{aligned} \rho u_{tt} &= F u_{xx}, \quad u(t, 0) = u(t, L) = 0, \\ u(0, x) &= f(x), \quad u_t(0, x) = g(x). \end{aligned}$$

If we look for time harmonic solutions $e^{i\omega t} v(x)$ then we are led to the eigenvalue problem

$$(2.16) \quad v_{xx} = -\lambda \rho(x) v, \quad v(0) = v(L) = 0, \quad \lambda = \frac{\omega^2}{F}.$$

Since $\rho(x) > 0$ is a random function, λ is a random variable, and the corresponding solution $v^{(\lambda)}(x)$ of (2.16) is a random function. Problems of this kind are also dif-

difficult to treat. Some results have been obtained in [1]. We shall not consider this problem here.

Finally we consider wave propagation in a one dimensional random medium. Let $u(t, x)$ represent the wave field at location x at time t . We shall assume that the index of refraction $n(x)$ is identically 1 outside the interval $[0, L]$ and equal to a random function inside $[0, L]$. Then $u(t, x)$ satisfies the problem

$$(2.17) \quad \begin{aligned} u_{tt} - \left(\frac{c}{n(x)} \right)^2 u_{xx} &= 0, \\ n(x) &= 1, \quad x < 0, \quad x > 0; \quad n(x) = 1 + \mu(x), \quad 0 \leq x \leq L, \\ u(0, x) &= f(x), \quad u_t(0, x) = g(x), \\ u(t, x), \quad u_x(t, x) &\text{ continuous at } x = 0 \text{ and } x = L. \end{aligned}$$

We assume that the random function $\mu(x)$, which has mean zero, is bounded in absolute value by a constant less than one, so that the above problem is well posed. Let us restrict our considerations to time harmonic fields of frequency ω by assuming that

$$(2.18) \quad u(t, x) = e^{i(\omega t - kx)} + R e^{i(\omega t + kx)}, \quad x < 0, \quad k = \omega/c,$$

$$(2.19) \quad u(t, x) = T e^{i(\omega t - kx)}, \quad x > L.$$

Here T and R are complex random variables and are called the **transmission** and **reflection coefficient** respectively. In view of (2.18), (2.19), problem (2.17) reduces to the following problem for $v(x) = e^{-i\omega t} u(t, x)$:

$$(2.20) \quad \begin{aligned} v_{xx} + k^2 n^2(x) v &= 0, \quad 0 \leq x \leq L \\ v(x) &= e^{ikx} + R e^{-ikx}, \quad x < 0, \\ v(x) &= T e^{ikx}, \quad x > L, \\ v, v_x &\text{ continuous at } x = 0 \text{ and } x = L. \end{aligned}$$

The physical significance of (2.18) and (2.19) is that, under steady state conditions, a plane wave $e^{i(\omega t - kx)}$, is incident upon the interval of inhomogeneities $[0, L]$ which produces a reflected wave $R e^{i(\omega t + kx)}$, $x < 0$, and a transmitted wave $T e^{i(\omega t - kx)}$, $x > L$. The problem we consider is to determine the statistical characteristics of T and R given the statistical characteristics of $n(x)$ in $0 \leq x \leq L$. In section 6 we shall give some results on this problem when case (ii) (see above) prevails. Note that here the space variable x plays the role of time, $v = 2\pi/k$ is the **wave length** of the incident wave, and l is the correlation length of the inhomogeneities $\mu(x)$, ($E\{\mu(x)\} = 0$, $E\{\mu(x)\mu(x')\} = \sigma^2 R(x - x')$).

3. Formulation of the mathematical problem. The description of the problems we

have given in section 2 and in the introduction has been somewhat loose. In this section we wish to state these problems in correct mathematical language. We shall do this for a general class of stochastic ordinary differential equations but we shall restrict ourselves to initial value problems. The other problems can be formulated similarly.

Let (Ω, \mathcal{F}, P) be a probability space, that is, Ω is an abstract set, \mathcal{F} a σ -algebra of subsets of Ω and P a probability measure on \mathcal{F} . Let $x(t, \omega)$, $\omega \in \Omega$, be a measurable mapping of $\Omega \times [0, \infty)$ to R^m (on $\Omega \times [0, \infty)$ we take the product measure $P \times \mu$, μ = Lebesgue measure). Let $F(z, x, t)$ be a mapping of $R^n \times R^m \times [0, \infty)$ into R^n such that it is differentiable in z and Lebesgue measurable in t and x . The differential equation

$$(3.1) \quad \frac{dz(t, \omega)}{dt} = F(z(t, \omega), x(t, \omega), t), \quad z(0, \omega) = z_0$$

has a unique solution $z(t, \omega)$ in the sense of Carathéodory [3]. The solution is a measurable function of ω . The measurability with respect to ω follows from considerations analogous to those yielding continuous dependence on parameters.

If the paths $x(t, \omega)$ are almost surely continuous and $F(z, x, t)$ is continuous as a function of x and t , then the paths $z(t, \omega)$ will be almost surely differentiable and will satisfy (3.1). The initial condition $z(0, \omega) = z_0$ may be a random variable without additional complication. Similarly, if $x(t, \omega)$ is a pure jump process with no instantaneous states and $F(z, x, t)$ is continuous as a function of x and t , then $z(t, \omega)$ will have continuous paths which are differentiable between jumps of $x(t, \omega)$.

Thus no question of "existence and uniqueness" arises for the stochastic differential equation (3.1) which requires considerations different from the deterministic situation. However, such questions do arise when the process $x(t, \omega)$ has unusually rough paths, as for example when $x(t, \omega)$ is white noise, i.e., the "derivative" of the Brownian motion process. Since the Brownian paths are almost surely non-differentiable an equation such as (3.1) must then be interpreted in an appropriate manner. One useful interpretation leads to the subject of Itô stochastic equations [16]. Under this interpretation the ensuing process $z(t, \omega)$ is a diffusion Markov process and the representation (3.1) can be exploited in the investigation of their properties [16].

From the point of view of applications, however, this interpretation is frequently inappropriate. The simplest reason why this is so is that the process is not associated with (3.1) in an invariant way. A change of coordinates in (3.1) leads to a different diffusion process (not the same process in the new coordinate representation). An interpretation which overcomes this difficulty, but which is not as convenient mathematically as Itô's, has been proposed by Stratonovich [25]. This point is also discussed in [27].

We shall always assume here that $x(t, \omega)$ has well-behaved paths so that no problem of interpretation arises. The main objective is to find effective methods for

obtaining the statistical properties of $z(t, \omega)$, such as $E\{z(t, \omega)\}$, etc., given those of $x(t, \omega)$. As is customary, we shall not write the $\omega \in \Omega$ explicitly in the sequel.

4. Equations with Markov coefficients. Let us consider the stochastic differential equation

$$(4.1) \quad \frac{dz(t)}{dt} = F(z(t), x(t)), \quad z(0) = z.$$

Here $F(z, x)$ is a mapping of $R^n \times R^m$ into R^n , differentiable in z and continuous in x , and $x(t)$ is an R^m valued Markov process. We first consider the case where $x(t)$ is a time homogeneous diffusion process whose infinitesimal mean and variance are given by

$$(4.2) \quad a_{ij}(x) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} E\{[x_i(t + \Delta t) - x_i(t)][x_j(t + \Delta t) - x_j(t)] \mid x(t) = x\}$$

$$(4.3) \quad b_j(x) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} E\{x_j(t + \Delta t) - x_j(t) \mid x(t) = x\} \quad i, j = 1, \dots, m.$$

From the above assumptions it follows that $(z(t), x(t))$ constitutes an R^{n+m} valued diffusion Markov process. To compute its infinitesimal mean and variance we use (4.1) and obtain, in addition to (4.2) and (4.3)

$$\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} E\{z_i(t + \Delta t) - z_i(t) \mid z(t), x(t)\} = F_i(z(t), x(t)), \quad i = 1, 2, \dots, n,$$

$$\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} E\{[z_i(t + \Delta t) - z_i(t)][x_j(t + \Delta t) - x_j(t)] \mid z(t), x(t)\} = 0, \\ i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m,$$

$$\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} E\{[z_i(t + \Delta t) - z_i(t)][z_j(t + \Delta t) - z_j(t)] \mid z(t), x(t)\} = 0, \\ i = 1, \dots, n, \quad j = 1, \dots, n.$$

Let $f(z, x)$ be a bounded smooth function on R^{n+m} and define $u(t, z, x)$ by

$$u(t, z, x) = E_{z, x}\{f(z(t), x(t))\}.$$

Here $E_{z, x}\{ \}$ denotes expectation with respect to the measure on the paths of the process $(z(t), x(t))$ starting from $(z, x) \in R^{n+m}$. Then $u(t, z, x)$ satisfies the backward Kolmogorov equation [7]

$$(4.4) \quad \frac{\partial u}{\partial t} = \frac{1}{2} \sum_{i, j=1}^m a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{j=1}^m b_j(x) \frac{\partial u}{\partial x_j} \\ + \sum_{j=1}^n F_j(z, x) \frac{\partial u}{\partial z_j}, \quad u(0, z, x) = f(z, x).$$

Similarly, it can be shown that the transition probability density of $(z(t), x(t))$, $P(t, z, x; z_0, x_0)$, when it exists, satisfies the forward equation [7]

$$(4.5) \quad \begin{aligned} \frac{\partial P}{\partial t} = & \frac{1}{2} \sum_{i,j=1}^m \frac{\partial^2}{\partial x_i \partial x_j} (a_{ij}(x)P) - \sum_{j=1}^m \frac{\partial}{\partial x_j} (b_j P) \\ & - \sum_{j=1}^n \frac{\partial}{\partial z_j} (F_j P), \end{aligned}$$

$$P(0, z, x; z_0, x_0) = \delta(x - x_0) \delta(z - z_0).$$

The coefficients a_{ij} , b_j and F_j are assumed to be sufficiently smooth.

The problem of determining the statistical characteristics of $z(t)$ defined by (4.1) is equivalent to solving (4.4) or (4.5) when $x(t)$ is a given diffusion process. For most problems of interest this is an untenable objective and approximations must be sought. We shall do this systematically in section 5 in a somewhat broader context. We shall explore here briefly the connections between stochastic processes and differential equations, a subject of independent interest. We shall also consider a simple example, for which (4.4) is not solvable explicitly but a time independent solution can be obtained.

Let $x(t)$ be a one dimensional diffusion process with diffusion coefficients $a(x)$ and drift coefficient $b(x)$. Let $z(t)$ be the scalar valued process defined by

$$(4.6) \quad \frac{dz(t)}{dt} = v(x(t))z(t), \quad z(0) = z.$$

Here $v(x)$ is a bounded smooth function on R . Clearly the solution of (4.6) is

$$(4.7) \quad z(t) = z \exp \left\{ \int_0^t v(x(s)) ds \right\}.$$

From the above considerations it follows that

$$(4.8) \quad u(t, z, x) = E_{z,x} \{ f(z(t), x(t)) \}$$

satisfies the problem

$$(4.9) \quad \begin{aligned} \frac{\partial u}{\partial t} = & \frac{1}{2} a(x) \frac{\partial^2 u}{\partial x^2} + b(x) \frac{\partial u}{\partial x} + v(x) z \frac{\partial u}{\partial z} \\ u(0, z, x) = & f(z, x). \end{aligned}$$

Consider the special initial function

$$f(z, x) = zg(x),$$

with $g(x)$ smooth. Assuming all integrals exist we define $V(x, t)$ by

$$(4.10) \quad V(x, t) = E_x \left[\exp \left\{ \int_0^t v(x(s)) ds \right\} g(x(t)) \right].$$

But from the definition of u it follows that

$$(4.11) \quad u(t, z, x) = z V(x, t).$$

Thus (4.11) and (4.9) yield

$$(4.12) \quad \frac{\partial V}{\partial t} = \frac{1}{2} a(x) \frac{\partial^2 V}{\partial x^2} + b(x) \frac{\partial V}{\partial x} + v(x)V, \quad V(x, 0) = g(x).$$

The representation (4.10) of the solution to (4.12) is called the Feynman-Kac formula [9]. The derivation we gave here is similar to that of Frisch [5]. Some examples of how this representation is exploited can be found in [6, 12].

Next we consider the case where $x(t)$ in (4.1) is not a diffusion process but a pure jump Markov process. For simplicity we take $x(t)$ to be a Markov chain with infinitesimal matrix Q .

$$Q = \left. \frac{d}{dt} e^{Qt} \right|_{t=0}.$$

Here e^{Qt} is the transition probability matrix of the chain. We assume that the chain has no instantaneous states

$$-q_{ii} < \infty, \quad i = 1, 2, \dots$$

and is conservative

$$-q_{ii} = \sum_{j \neq i} q_{ij}.$$

Again $(z(t), x(t))$ are jointly a Markov process with state space $R^n \times \{1, 2, \dots\}$. Since $x(t)$ is discrete valued we shall denote functional dependence on $x(t)$ by a subscript. Thus, as above, the functions

$$(4.13) \quad u_i(t, z) = E_{i,z} \{ f_{x(t)}(z(t)) \},$$

with $f_i(z)$, $i = 1, 2, \dots$, smooth functions on R^n , satisfy the system of partial differential equations

$$(4.14) \quad \begin{aligned} \frac{\partial u_i}{\partial t} &= \sum_{j=1}^{\infty} Q_{ij} u_j + \sum_{k=1}^n F_k(z, x_i) \frac{\partial u_i}{\partial z_k}, \\ u_i(0, z) &= f_i(z). \end{aligned}$$

Let us consider the special case where $x(t)$ is the random telegraph process, i.e., the two state chain with values $+1$ and -1 and

$$(4.15) \quad Q = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}, \quad a > 0.$$

Let us further define F to be scalar valued, so that $z(t)$ is scalar valued, and set

$$(4.16) \quad F(z, \pm 1) = \pm 1.$$

Then from (4.15) and (4.16) it follows that for this example (4.14) becomes

$$(4.17) \quad \begin{aligned} \frac{\partial u_+}{\partial t} &= -au_+ + au_- + \frac{\partial u_+}{\partial z}, \\ \frac{\partial u_-}{\partial t} &= au_+ - au_- - \frac{\partial u_-}{\partial z}, \quad u_{\pm}(0, z) = f_{\pm}(z). \end{aligned}$$

In (4.17) we have denoted by \pm the two states of the chain instead of numerical subscripts. Now from (4.16) and (4.1) it follows that

$$z(t) = z + \int_0^t x(s) ds,$$

and from (4.13)

$$(4.18) \quad u_{\pm}(t, z) = E_{\pm} \left\{ f_{x(t)} \left(z + \int_0^t x(s) ds \right) \right\}.$$

This result is a probabilistic representation of the solution of the hyperbolic system (4.17) in the form of an expectation over paths of the random telegraph process. It was first obtained by Kac [11]. However, the connection of this representation to the Feynman-Kac formula (4.10) and (4.12) was not noticed there. The probabilistic representation of the solution of hyperbolic systems is given in [8, 23].

As an example of the usefulness of the representation (4.18) we shall employ it to obtain information about properties of the solutions of (4.17). We do this in somewhat greater generality. Let $x(t)$ be an N -state ergodic Markov chain with values x_1, x_2, \dots, x_N , and with a unique invariant probability vector (P_1, P_2, \dots, P_N) so that

$$(4.19) \quad \sum_{i=1}^N P_i x_i = 0.$$

We assume that $z(t)$ is scalar valued and

$$(4.20) \quad F(z, x_i) = x_i, \quad i = 1, \dots, N.$$

Then the functions $u_i(t, z) = E_i \{ f_{x(t)}(z + \int_0^t x(s) ds) \}$ satisfy the differential equations

$$(4.21) \quad \frac{\partial u_i}{\partial t} = \sum_{j=1}^N Q_{ij} u_j + x_i \frac{\partial u_i}{\partial z}, \quad u_i(0, z) = f_i(z).$$

We wish to study the behavior of $u_i(t, z)$ when t is large. For this purpose, we introduce a small parameter ε which we allow to tend to zero as $t \rightarrow \infty$. In this way we obtain a nontrivial limit. We set

$$(4.22) \quad u_i^{(\varepsilon)}(\tau, z) = E_i \left\{ f_{x(\tau/\varepsilon^2)} \left(z + \varepsilon \int_0^{\tau/\varepsilon^2} x(s) ds \right) \right\}, \quad \tau = \varepsilon^2 t.$$

Thus $u_i^{(\varepsilon)}$ satisfies the system

$$(4.23) \quad \frac{\partial u_i^{(\varepsilon)}}{\partial \tau} = \frac{1}{\varepsilon^2} \sum_{j=1}^N Q_{ij} u_j^{(\varepsilon)} + \frac{1}{\varepsilon} x_i \frac{\partial u_i^{(\varepsilon)}}{\partial z},$$

$$u_i(0, z) = f_i(z).$$

We note that when we let $\varepsilon \rightarrow 0$ formally in (4.23), $0 < \tau \leq \tau_0$ fixed, it reduces to an algebraic problem and thus it is a "singular" perturbation. It is not difficult however to find the limit of $u_i^{(\varepsilon)}(\tau, z)$ by using (4.22). Note that

$$u_i^{(\varepsilon)}(\tau, z) = E_i \left\{ f_{x(\tau/\varepsilon^2)} \left(z + \sqrt{\tau} \frac{1}{\sqrt{t}} \int_0^t x(s) ds \right) \right\}.$$

Now $(1/\sqrt{t}) \int_0^t x(s) ds$ converges weakly as $t \rightarrow \infty$ to a Gaussian random variable with mean zero and variance

$$\sigma^2 = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t \int_0^t E\{x(s)x(s')\} ds ds'.$$

This is the Central Limit Theorem for Markov Chains [17]. Moreover

$$\lim_{\substack{\varepsilon \rightarrow 0 \\ 0 < \tau \leq \tau_0}} P\{x(\tau/\varepsilon^2) = x_k\} = P_k$$

independently of the initial conditions on $x(t)$. This follows from the ergodic theorem for Markov chains [17]. These two observations imply that when $f_i(z)$, $i = 1, \dots, N$, are bounded continuous functions, then

$$\lim_{\substack{\varepsilon \rightarrow 0 \\ 0 < \tau \leq \tau_0}} u_i^{(\varepsilon)}(\tau, z) = \int_{-\infty}^{\infty} \left\{ \sum_{i=1}^N P_i f_i(z + \sqrt{\tau} \zeta) \right\} \frac{e^{-\zeta^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} d\zeta.$$

We summarize the above as follows. Under the hypothesis stated, the solution of (4.23) converges as $\varepsilon \rightarrow 0$, $0 < \tau \leq \tau_0$ fixed, uniformly in z to $u^{(0)}(\tau, z)$ which is independent of i and satisfies the problem

$$\frac{\partial u^0}{\partial \tau} = \frac{1}{2} \sigma^2 \frac{\partial^2 u^0}{\partial z^2}, \quad u^0(0, z) = \sum_{i=1}^N P_i f_i(z).$$

The use of differential equations methods to prove limit theorems goes back to Khinchine [13]. It was used by Pinsky [23] to investigate (4.23) and obtain limit theorems for Markov chains with error estimates. Here we have used the limit theorems to obtain a result on the singular perturbation of (4.21) (in the form (4.23)) as was done in [8].

The relation between limit theorems for Markov processes and singular perturbation of differential equations has received considerable attention recently. The article of Pinsky [24] gives a good survey of the subject.

We close this section with an example where a time independent solution of (4.4)

can be obtained. More complicated examples and some remarks on the existence of time independent (or equilibrium) solutions for diffusion equations can be found in [26].

Consider the nonlinear oscillator

$$(4.24) \quad \frac{dz_1^2}{dt^2} + k^2 z_1 + a^2 z_1^3 = N(t), \quad z_1(0) = z_1, \quad \frac{dz_1(0)}{dt} = z_2.$$

Here a and k are real constants and $N(t)$ is the white noise process, that is, the formal derivative of the Brownian motion process. No question of interpretation for (4.24) arises (see Sec. 3) because we may simply integrate (4.24) once and consider the resulting equation as defining $z_1(t)$. Let

$$z_2(t) = \frac{dz_1(t)}{dt}.$$

Because of the properties of Brownian motion (it has independent increments) the process $(z_1(t), z_2(t))$ is a diffusion Markov process with backward Kolmogorov equation [7]

$$(4.25) \quad \frac{\partial u}{\partial t} = \frac{1}{2} \frac{\partial^2 u}{\partial z_2^2} + z_2 \frac{\partial u}{\partial z_1} - (k^2 z_1 + a^2 z_1^3) \frac{\partial u}{\partial z_2},$$

$$u(0, z_1, z_2) = f(z_1, z_2).$$

Let us define the function $H(z_1, z_2)$ by

$$H(z_1, z_2) = \frac{1}{2} k^2 z_1^2 + \frac{1}{4} a^2 z_1^4 + \frac{1}{2} z_2^2.$$

This is the total energy, or Hamiltonian, of the oscillator. We note that

$$u(z_1, z_2) = c e^{-H(z_1, z_2)}$$

is a time independent solution of (4.25). The constant c may be chosen so that $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u(z_1, z_2) dz_1 dz_2 = 1$. Then $u(z_1, z_2)$ is an invariant or equilibrium density for the process $(z_1(t), z_2(t))$. It is called the Maxwell-Gibbs distribution of the oscillator (4.24). The remarkable fact is that although for (4.24) we cannot solve the deterministic nonlinear problem, we have a simple description of the equilibrium behavior of the stochastic problem.

5. The Theorem of Stratonovich and Has'minskii. In the previous section, we saw that when the coefficients of the stochastic equation are Markovian it is still practically impossible to solve the Kolmogorov equations. Sometimes it may be possible to obtain time independent solutions. When we do not assume that the coefficients of the equation are Markovian, the situation is worse since the apparatus of differential equations is not available. We shall now consider the main objective

of our study of stochastic equations: the construction of effective approximation methods. Before we can state results in this direction we must first clarify the sense in which approximations will be considered.

Let us return to the harmonic oscillator example of section 2. We rewrite it here in a slightly different notation:

$$(5.1) \quad m \frac{d^2 u}{dt'^2} + k'(t')u = 0, \quad u(0) = u_0, \quad \frac{du}{dt'}(0) = \dot{u}_0.$$

We assume that $k'(t')$ is a stationary random function with mean k'_0 and write

$$k'(t') = k'_0 [1 + \sigma x'(t')].$$

Thus

$$E\{x'(t')\} = 0, \quad E\{x'(t')x'(s')\} = R'(t' - s'),$$

$$l = \int_0^\infty R'(s) ds < \infty.$$

The dimensionless parameter σ^2 is the variance of the fluctuations of $x'(t')$ and $R'(s')$ is its correlation function. We have also assumed that the correlation time l is finite. Let us define a dimensionless time scale

$$t = t' / l.$$

Then u as a function of t satisfies the stochastic equation

$$(5.2) \quad \frac{d^2 u}{dt^2} + k_0 u + \varepsilon x(t)u = 0, \quad u(0) = u_0, \quad \frac{du}{dt}(0) = l\dot{u}_0.$$

Here we have introduced the notation

$$x(t) = x'(tl), \quad k_0 = \frac{k'_0 l^2}{m}, \quad \varepsilon = k_0 \sigma.$$

In system form (5.2) becomes

$$(5.3) \quad \frac{d}{dt} \begin{pmatrix} u \\ v \end{pmatrix} = A \begin{pmatrix} u \\ v \end{pmatrix} + \varepsilon X(t) \begin{pmatrix} u \\ v \end{pmatrix}, \quad \begin{pmatrix} u \\ v \end{pmatrix}(0) = \begin{pmatrix} u_0 \\ l\dot{u}_0 \end{pmatrix}$$

$$A = \begin{pmatrix} 0 & 1 \\ -k_0 & 0 \end{pmatrix}, \quad X(t) = \begin{pmatrix} 0 & 0 \\ -x(t) & 0 \end{pmatrix}.$$

The role of ε in the dimensionless formulation of the oscillator problem is now clear. We shall assume $\varepsilon \ll 1$ and allow t to be large. Since $E\{X(t)\} = 0$, stochastic effects will become important for times t of the order $1/\varepsilon^2$ (in the t units the correlation time is 1). This corresponds to case (ii) of section 2. Therefore, straightforward perturbation expansion in power series of ε and then averaging of the result will not yield interesting results since they are valid only for short times. We shall

state below a limit theorem due to Has'minskii [9] and Stratonovich [25] which characterizes the behavior of the stochastic process defined by (5.3) in the limit $\varepsilon \rightarrow 0$, $\tau \rightarrow \infty$, $\varepsilon^2 t = \text{constant}$.

The result is actually more general. It applies to a class of stochastic equations which we shall describe. Before doing this, however, we shall motivate the form of the general problem (see (5.4) below) by reconsidering (5.3).

Let $z(t) = \begin{pmatrix} z_1(t) \\ z_2(t) \end{pmatrix}$ be defined by

$$z(t) = e^{-At} \begin{pmatrix} u(t) \\ v(t) \end{pmatrix}.$$

From (5.3) we find that the vector-valued process $z(t)$ satisfies the equation

$$\frac{dz(t)}{dt} = \varepsilon [e^{-At} X(t) e^{At}] z(t), \quad z(0) = z_0 = \begin{pmatrix} u_0 \\ l\dot{u}_0 \end{pmatrix}.$$

We shall call $z(t)$ the **slowly varying part** of the process $\begin{pmatrix} u \\ v \end{pmatrix} = e^{At} z(t)$ because the rate of change of z is of order ε . Our main interest is the characterization of the limit of the process $z(t)$, which depends on ε , as ε tends to zero, t tends to infinity and $\varepsilon^2 t$ remains fixed. We now state the Stratonovich-Has'minskii theorem which provides a very satisfactory answer to this question for a large class of problems.

Consider the stochastic process $z(t)$ with values in R^n defined by the equation

$$(5.4) \quad \frac{dz(t)}{dt} = \varepsilon F(z(t), x(t), t), \quad z(0) = z_0.$$

Here $x(t)$ is an R^m valued stochastic process and F is a mapping of $R^n \times R^m \times [0, \infty)$ into R^n such that for some constant C

$$(5.5) \quad |F_i| < C, \quad \left| \frac{\partial F_i}{\partial z_j} \right| < C, \quad \left| \frac{\partial F_i}{\partial z_j \partial z_k} \right| < C, \quad i, j, k = 1, \dots, n,$$

uniformly in z , x and t . F is a measurable function of x and t for fixed z . Assume that

$$(5.6) \quad E\{F(z, x(t), t)\} = 0$$

and the limits

$$(5.7) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \int_{t_0}^{t_0+T} \int_{t_0}^s E \left\{ \sum_{j=1}^n \frac{\partial F_i(z, x(s), s)}{\partial z_j} F_j(z, x(\sigma), \sigma) \right\} d\sigma ds = b_i(z),$$

$$(5.8) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \int_{t_0}^{t_0+T} \int_{t_0}^s E\{F_i(z, x(s), s) F_j(z, x(\sigma), \sigma)\} d\sigma ds = a_{ij}(z),$$

exist uniformly in t_0 and z . Assume further that if U_s^t , $0 \leq s \leq t \leq \infty$, denotes the σ -algebra of events generated by $x(\sigma)$, $s \leq \sigma \leq t$, then

$$(5.9) \quad \sup_{\substack{A \in \mathcal{U}_0 \\ B \in \mathcal{U}_{t+s}}} |P(B|A) - P(B)| < \beta(s),$$

and $\beta(s) \downarrow 0$ as $s \rightarrow \infty$ so that $s^6 \beta(s) \downarrow 0$ as well. Under the above assumptions the process $z^{(\varepsilon)}(\tau)$ defined by

$$(5.10) \quad z^{(\varepsilon)}(\tau) = z(\tau/\varepsilon^2), \quad \tau = \varepsilon^2 t,$$

converges weakly as $\varepsilon \rightarrow 0$, $0 \leq \tau \leq \tau_0$ to a Markov process $z^0(\tau)$ which is continuous with probability 1 and whose infinitesimal generator is given by

$$(5.11) \quad L = \sum_{i,j=1}^n a_{ij}(z) \frac{\partial^2}{\partial z_i \partial z_j} + \sum_{j=1}^n b_j(z) \frac{\partial}{\partial z_j}.$$

This remarkable result was enunciated by Stratonovich [26] who applied it successfully to a number of problems. A somewhat different form of the above formulation of the theorem, its proof, and several interesting remarks can be found in [25]. Here we shall present an application of the result to partial differential equations using the representations of section 4.

Suppose that $x(t)$ in (5.4) is an R^m valued diffusion Markov process which has an everywhere positive transition probability density $P(t; x, x_0)$ and a unique positive invariant measure $p(dx)$. Suppose all conditions of Has'minskii theorem are satisfied. For any bounded continuous function on R^{n+m} the function

$$(5.12) \quad u(s, t; z, x) = E_{z,x;s} \{f(z(t), x(t))\}$$

satisfies the backward Kolmogorov equation

$$-\frac{\partial u}{\partial s} = \frac{1}{2} \sum_{i,j=1}^m A_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{j=1}^m B_j(x) \frac{\partial u}{\partial x_j} + \varepsilon \sum_{j=1}^n F_j(z, x, s) \frac{\partial u}{\partial z_j},$$

$$s < t, \quad u(t, t; z, x) = f(z, x).$$

Here $A_{ij}(x)$ and $B_j(x)$ are the infinitesimal variance matrix and drift vector of the process $x(t)$ and $E_{x,z;s}\{\}$ denotes expectation over the paths of the Markov process $(z(t), x(t))$ starting from (z, x) when $t = s$. Let

$$\sigma = \varepsilon^2 s, \quad \tau = \varepsilon^2 t, \quad u^{(\varepsilon)}(\sigma, \tau, z, x) = u(\sigma/\varepsilon^2, \tau/\varepsilon^2, z, x).$$

Then,

$$-\frac{\partial u^{(\varepsilon)}}{\partial \sigma} = \frac{1}{2} \frac{1}{\varepsilon^2} \sum_{j,i=1}^m A_{ij}(x) \frac{\partial^2 u^{(\varepsilon)}}{\partial x_i \partial x_j} + \frac{1}{\varepsilon^2} \sum_{j=1}^m B_j(x) \frac{\partial u^{(\varepsilon)}}{\partial x_j}$$

$$+ \frac{1}{\varepsilon} \sum_{j=1}^n F_j(z, x, \sigma/\varepsilon^2) \frac{\partial u^{(\varepsilon)}}{\partial z_j},$$

$$\sigma < \tau, \quad u^{(\varepsilon)}(\tau, \tau; z, x) = f(z, x).$$

It follows immediately from the Stratonovich-Hashminskii theorem and (5.12) that $u^{(e)}(0, \tau; z, x)$ converges uniformly in z , $0 < \tau \leq \tau_0$, to $u^{(0)}(\tau, z)$, independently of x where $u^{(0)}$ satisfies the problem

$$\frac{\partial u^{(0)}}{\partial \tau} = Lu^{(0)}, \quad u^{(0)}(0, z) = \int f(z, x)p(dx).$$

The correct initial conditions for $u^{(0)}$ are obtained in a manner similar to that used in obtaining the result for (4.23). The coefficients $a_{ij}(z)$, $b_j(z)$ in the operator L are obtained from (5.7) and (5.8) where $E\{\cdot\}$ denotes expectation over the paths of $x(t)$ starting at $t = 0$ with the invariant measure $p(dx)$. Thus in (5.7) and (5.8) the expectation can be written explicitly as integration with respect to appropriate weights involving $p(dx)$ and $P(t, x, x_0)$. Let us note that the singular perturbation result we have just deduced, even formally, is a nontrivial result to obtain.

When $x(t)$ is a Markov chain we can again apply the Stratonovich-Has'minskii theorem and obtain a singular perturbation result similar to that of (4.23) where a hyperbolic system converges to a diffusion equation. Now, however, we may allow F to depend on z, x, t in any manner compatible with the hypothesis of the theorem stated above.

The significance of the Stratonovich-Has'minskii result can best be understood when it is seen as a generalization of the classical central limit theorem to processes that are functionals of asymptotically independent processes (i.e., condition (5.9) holds) defined via a differential equation. Indeed if $F(z, x, t) \equiv x$ then we trivially obtain the classical central limit theorem [4]. The usefulness of this generalized central limit theorem is, in practice, limited severely by the fact that few diffusion equations with variable coefficients are solvable explicitly. In any case, this is the best that can be achieved under the circumstances.

6. Examples. In this section we shall apply the result of Has'minskii and Stratonovich to two problems. One is the harmonic oscillator problem (5.3) and the other is the wave propagation problem (2.20).

Let us transform (5.3) by introducing slowly varying dependent variables. We note that $F(z, x, t)$ is linear in z here and thus (5.5) is violated. To circumvent this difficulty we change coordinates appropriately. Let us write (5.3) in slowly varying form:

$$\begin{aligned} \dot{z}_1 &= \varepsilon x(t) \left[\frac{z_1}{\sqrt{k_0}} \sin(\sqrt{k_0}t) \cos(\sqrt{k_0}t) + \frac{z_2}{k_0} \sin^2(\sqrt{k_0}t) \right], \quad z_1(0) = z_1 \\ \dot{z}_2 &= \varepsilon x(t) \left[-z_1 \cos^2(\sqrt{k_0}t) - \frac{z_2}{\sqrt{k_0}} \sin(\sqrt{k_0}t) \cos(\sqrt{k_0}t) \right], \quad z_2(0) = z_2. \end{aligned}$$

The change of variables

$$z_1 = e^r \cos \theta, \quad z_2 = -\sqrt{k_0} e^r \sin \theta, \quad -\infty < r < \infty, \quad 0 \leq \theta \leq 2\pi,$$

leads to a system of equations of the form

$$(6.1) \quad \begin{aligned} \dot{r} &= \varepsilon x(t) g_1(\theta, t), & r(0) &= r_0, \\ \dot{\theta} &= \varepsilon x(t) g_2(\theta, t), & \theta(0) &= \theta_0. \end{aligned}$$

The functions g_1 and g_2 are trigonometric polynomials in θ and t and thus in the representation (6.1) the oscillator problem satisfies (5.5) provided $x(t)$ is a bounded random process, which we now assume. We assume further that $x(t)$ is stationary with mean zero, has correlation function $E\{x(t+s)x(t)\} = R(s)$, and satisfies (5.9). Thus all hypotheses are satisfied since explicit computation shows that the limits (5.7) and (5.8) exist uniformly in t_0 and are independent of r and θ . It follows therefore that L , defined by (5.11), has constant coefficients and is given by

$$\begin{aligned} L &= \frac{b}{2} \frac{\partial^2}{\partial r^2} + b \frac{\partial}{\partial r} + \left(a + \frac{b}{2}\right) \frac{\partial^2}{\partial \theta^2} + c \frac{\partial}{\partial \theta}, \\ a &= k_0 \frac{S(0)}{4}, \quad b = k_0 \operatorname{Re} \frac{S(2\sqrt{k_0})}{4}, \quad c = k_0 \operatorname{Im} \frac{S(2\sqrt{k_0})}{4}, \\ S(\omega) &= \int_0^\infty R(\sigma) e^{i\omega\sigma} d\sigma. \end{aligned}$$

The statistical properties of the limit process $z^{(0)}(\tau)$ can be obtained from those of $r^{(0)}(\tau)$ and $\theta^{(0)}(\tau)$ whose transition probability density is the solution of the initial value problem

$$\partial P / \partial \tau = LP, \quad P(0, r, \theta; r_0, \theta_0) = \delta(r - r_0) \delta(\theta - \theta_0).$$

This equation can be solved explicitly since it has constant coefficients. Then by transforming to the original variables in (5.3) we find that

$$(6.2) \quad \begin{aligned} E\{u(t)\} &= \exp\{k_0/4 [\operatorname{Re} S(2\sqrt{k_0}) - S(0)] \varepsilon^2 t\} \\ &\quad \cos \left(\sqrt{k_0} - \frac{\varepsilon^2 k_0 \operatorname{Im} S(2\sqrt{k_0})}{4} \right) t + o(1), \\ &\quad 0 \leq t \leq 1/\varepsilon^2. \end{aligned}$$

An interesting feature of (6.2) is that the random fluctuations in the spring constant cause the mean displacement to decay with time on the slow time scale $\varepsilon^2 t$. Similarly we find that

$$(6.3) \quad \begin{aligned} E\{u^2(t)\} &= \frac{1}{2} \exp \left\{ \frac{k_0}{2} [\operatorname{Re} S(2\sqrt{k_0}) - 2S(0)] \varepsilon^2 t \right\} \cos \left(2\sqrt{k_0} - \frac{\varepsilon^2 k_0}{2} \operatorname{Im} S(2\sqrt{k_0}) \right) t \\ &\quad + \frac{1}{2} \exp(k_0 \operatorname{Re} S(2\sqrt{k_0}) \varepsilon^2 t) + o(1), \quad 0 \leq t \leq 1/\varepsilon^2. \end{aligned}$$

The result (6.3) shows that the mean square displacement is growing on the time

scale $\varepsilon^2 t$. This follows from the fact that $\text{Re} S(\omega)$ is the Fourier cosine transform of a correlation function and hence is positive.

Let us now consider problem (2.20). We assume that

$$n^2(x) = 1 + \varepsilon z(x),$$

where x is the "time" variable and $z(x)$ is a bounded zero mean stationary stochastic process for which (5.9) holds.

We wish to find $E\{|T|^2\}$, the mean square of the transmission coefficient, when L , the width of the region of the inhomogeneities, is of order $1/\varepsilon^2$. One way of treating (2.20) is presented in [18], [21]. Another way [20] is to find a stochastic equation for $T = T(L)$ as a function of the width L and apply the theorem of section 5 to it. We omit the calculations here and state the result:

$$E\{|T(L)|^2\} = \frac{4}{\sqrt{\pi}} e^{-\varepsilon^2 s L/4} \int_0^\infty \frac{x^2 e^{-x^2} dx}{\cosh(\varepsilon \sqrt{s L} x)} + o(1),$$

$$0 \leq L \leq 1/\varepsilon^2,$$

$$s = \frac{k^2}{2} \int_0^\infty R(t) \cos(2k\tau) d\tau, \quad R(\tau) = E\{z(x+\tau)z(x)\}.$$

From this result we find that the mean power transmission coefficient is approximately $1/2$ when the dimensionless quantity $\varepsilon^2 s L$ is of order 1.

Some other applications of an operator theoretic version of this result [22] are given in [2]. Other interesting applications are considered in [19].

Acknowledgment. The author wishes to thank the editor, and Professor J. B. Keller, for suggesting several improvements in the manuscript.

References

1. W. E. Boyce, Random eigenvalue problems, in Probabilistic Methods in Applied Mathematics, A. T. Bharucha-Reid, editor, Academic Press, New York, 1968.
2. R. Burridge, and G. C. Papanicolaou, The geometry of coupled mode propagation in random media, *Comm. Pure Appl. Math.*, to appear.
3. E. A. Coddington, and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, 1955.
4. W. Feller, *An Introduction to Probability Theory and Its Applications*, Vols. I, II, Wiley, New York, 1968.
5. U. Frisch, Wave propagation in random media, in Probabilistic Methods in Applied Mathematics, A. T. Bharucha-Reid, editor, Academic Press, New York, 1968.
6. I. M. Gelfand, and A. M. Yaglom, Integration in functional spaces and its applications to quantum physics, *J. Math. Phys.*, 1 (1960) 48-69.
7. I. I. Gikhman and A. V. Skorokhod, *Introduction to the Theory of Random Processes*, Saunders, Philadelphia, 1969.
8. R. Griego and R. Hersh, Theory of random evolution with applications to partial differential equations, *Trans. Amer. Math. Soc.*, 156 (1971) 405-418.
9. R. Z. Has'minskii, A limit theorem for the solution of differential equations with random right-hand sides, *Theory of Prob. and Applications*, 11 (1966) 390-406.

10. M. Kac, On the distribution of certain Wiener functionals, *Trans. Amer. Math. Soc.*, 65 (1949) 1–13.
11. ———, Some Stochastic Problems in Physics and Mathematics, Magnolia Petroleum Co. Lectures in Pure and Applied Science, No. 2, 1956.
12. ———, Probability and Related Topics in the Physical Sciences, Interscience, New York, 1959.
13. A. I. Khinchine, *Asymptotische Gesetze der Wahrscheinlichkeitsrechnung*, Chelsea, New York, 1948.
14. R. Kubo, Stochastic Liouville equation, *J. Math. Phys.*, 4 (1963) 174–183.
15. M. Lax, Classical Noise IV; Langevin Methods, *Rev. Mod. Phys.*, 38 (1966) 561–566.
16. H. P. McKean Jr., *Stochastic Integrals*, Academic Press, New York, 1969.
17. M. Loève, *Probability Theory*, Van Nostrand, Princeton, N. J., 1963.
18. J. A. Morrison, Application of a limit theorem to solutions of a stochastic differential equation, *J. Math. Anal. Appl.*, 39 (1972) 13–35.
19. J. A. Morrison and J. McKenna, article in *Proceedings of Symposium on Stochastic Equations*, SIAM-AMS, vol. 6, to appear.
20. G. C. Papanicolaou, Wave propagation in a one-dimensional random medium, *SIAM J. on Appl. Math.*, 21 (1971) 13–18.
21. ———, and J. B. Keller, Stochastic differential equations with applications to random harmonic oscillators and wave propagation in random media, *SIAM J. on Appl. Math.*, 21 (1971) 287–305.
22. ———, and R. Hersh, Some limit theorems for stochastic equations and applications, *Indiana Univ. Math. J.*, 21 (1972) 815–840.
23. M. Pinsky, Differential equations with a small parameter and the central limit theorem for functions on a finite Markov chain, *J. Wahrscheinlichkeitstheorie Vern. Gebiete*, 9 (1968) 101–111.
24. ———, Multiplicative operator functionals and their asymptotic properties, in *Advances in Probability*, vol. 3, Marcel Dekker, New York.
25. R. L. Stratonovich, *Conditional Markov Processes and Their Application to the Theory of Optimal Control*, Elsevier, New York, 1968.
26. ———, *Topics in the Theory of Random Noise*, Vol. I, II, Gordon and Breach, New York, 1963.
27. W. M. Wonham, Random differential equations in control theory, in *Probabilistic Methods in Applied Mathematics*, Vol. 2, A. T. Bharucha Reid, editor, Academic Press, New York, 1970.

THE MAIN CRISES

S. BIRNBAUM, Bronx Community College

In the November 1971 issue of this magazine, Gail S. Young discussed some crises of the mathematical world in seven parts. I think the crises listed in sixth and seventh place such as the Viet Nam war and unemployment should be in first place. Furthermore, I must disagree with a statement such as this: "Some of the problems — for example the war in Viet Nam — are ones that we as mathematicians, or the organization we represent, can do nothing much about."

This article does not represent an official position of either the Editors of the *American Mathematical Monthly*, or the *Mathematical Association of America*. Approximately once a month we receive a suggestion for starting a new section of the *Monthly*, and fortunately or unfortunately, we cannot accept almost all such suggestions. *Editor.*

The employment crisis is mentioned as "another conceivably controversial example." I'm sure that young Ph.D.'s who have failed to obtain employment would want this treated as problem number one.

I think that we as mathematicians and citizens must come to effective grips with the above-mentioned and other problems which come under the heading of socio-economic-political issues. It is very comforting to believe that mathematics could make great contributions in helping society solve its problems — as goes mathematics so goes society. But it is even truer that as goes society so goes mathematics. Thus the mere existence of problems requiring the participation of mathematicians for their solution does not guarantee full employment for mathematicians. We don't need any America and Africa that have the gravest problems in feeding their populations employ the least numbers of mathematicians.

I am one of those optimists who believe that the more mathematicians employed for peaceful purposes in any country, the better off it must be. From this it follows that I, like most mathematicians, would prefer to see increased production and employment of mathematicians. If such is to the interest of our country and its people, then it is our duty to tackle the problem of unemployment or underemployment and help solve it for the general good. Indeed, failure to work actively for full employment may result in loss for not a few mathematicians who fancy themselves in secure positions.

Two years ago I wrote a little article about the spectre of unemployment which would increasingly haunt us mathematicians among others. In it I predicted that our economy was approaching a stage of chronic crisis. When I showed it to some of my colleagues and asked if I should submit it for publication they told me, "Go ahead but it won't be printed!"

This is the attitude we must put behind us. We must try to effect the development of this country in a peaceful and progressive direction. We must agree that this country has many problems requiring the participation of mathematicians in their solution; but we must also recognize that this is no guarantee of full employment. If we agree with all this then we enter the frontier of the most controversial part of this analysis: Concretely, how can we go about helping to direct this country onto a path that is best for its people and thus its mathematicians? At this point, a statement of a concrete program for action in the MONTHLY would be in opposition to the tradition of avoidance of controversial issues. Therefore the first step must consist in realizing that the responsibility of the mathematical community in influencing our country in this time of growing crisis takes precedence over tradition. We must be willing to enter the area of dialogue over controversial issues. To this end we can make a modest beginning by reserving, in the MONTHLY, a small section to be devoted to discussing the socio-economic-political problems 'bugging' our members.

To the younger and the impatient members this may be a mouse of a conclusion for a mountainous beginning; but it is a first step and as such, small as it is, is the most important one.

MATHEMATICAL NOTES

EDITED BY ROBERT GILMER

The present backlog for this Department is substantial. Until further notice, new manuscripts cannot be accepted. This moratorium will probably continue until June 1, 1973; authors are requested to hold their manuscripts pending a further announcement.

THE SIGN OF THE BERNOULLI NUMBERS

L. J. MORDELL[†]

The **Bernoulli numbers** are defined by the usual expansion,

$$(1) \quad \frac{x}{e^x - 1} = 1 - \frac{x}{2} + \sum_{n=1}^{\infty} (-1)^{n-1} B_n x^{2n} / (2n)! = \sum_{n=0}^{\infty} \frac{b_n x^n}{n!},$$

say. It is classic that $B_n > 0$.

A proof by Euler follows from his formula,

$$(2n + 1) B_n = \sum_{r=1}^{n-1} \binom{2n}{2r} B_r B_{n-r},$$

for which an involved proof is given in Nielsen [1]. Well known are the analytic proofs leading to

$$B_r = \frac{(2r)!}{2^{2r-1} \pi^{2r}} \sum_{n=1}^{\infty} \frac{1}{n^{2r}}.$$

A proof using Bernoulli polynomials is given in Nörlund [2]. Lastly, an arithmetic proof is given by Uspensky and Heaslet [3], which is rather complicated and unilluminating.

I am not aware of any proof based upon the simplest principles and so the one now given may be of interest. We have

$$\frac{x}{e^x + 1} = \frac{x}{e^x - 1} - \frac{2x}{e^{2x} - 1} = \sum_{r=0}^{\infty} (1 - 2^r) \frac{b_r x^r}{r!}.$$

Multiply both sides by $x/(e^x - 1)$ and substitute from (1) the expansion for $2x/(e^{2x} - 1)$. Then

$$\frac{x}{2} \sum_{n=0}^{\infty} \frac{b_n 2^n x^n}{n!} = \sum_{r=0}^{\infty} (1 - 2^r) \frac{b_r x^r}{r!} \sum_{s=0}^{\infty} \frac{b_s x^s}{s!}.$$

Equate coefficients of x^{2n} on both sides. Since b_n is zero if n is > 1 and odd, we have

[†] Deceased March 11, 1972.

$$\sum_{r=0}^{\infty} (1-2^r) \frac{b_r}{r!} \cdot \sum_{s=0}^{\infty} \frac{b_s}{s!} = 0, \quad (r+s=2n, \quad r>0, \quad s>0).$$

The terms with $r=0,1$ contribute nothing to the left hand side. Isolating the term with $r=2n, s=0$, we have

$$(1-2^{2n}) \frac{b_{2n}}{(2n)!} + \sum_{r,s} \frac{(1-2^r)b_r b_s}{r!s!} = 0, \quad (r+s=2n, rs \neq 0),$$

where we may suppose that r and s are both even. We assume now that $(-1)^{m-1}b_{2m} > 0$ for $1 \leq m \leq n-1$. This is true for $m=1$, and we prove that it holds for $m=n$. The term in the summation has the sign

$$(-1)^{1+(r/2)-1+(s/2)-1} = (-1)^{n-1}. \text{ Hence } (-1)^{n-1}b_{2n} > 0,$$

and this finishes the proof.

References

1. N. Nielsen, *Traité élémentaire des nombres de Bernouilli*, Gauthier Villars, Paris, 1923, p.42.
2. N. H. Nörlund, *Vorlesungen über Differenzenrechnung*, Chelsea, New York, 1924, pp. 22-23.
3. J. V. Uspensky and M. A. Heaslet, *Elementary Number Theory*, McGraw-Hill, New York, 1939, Chapter IX.

THE SIGN OF THE BERNOULLI AND EULER NUMBERS

LEONARD CARLITZ and RICHARD SCOVILLE, Duke University

Put $x \cot x = \sum_{n=0}^{\infty} (-1)^n B_{2n} x^{2n}/(2n)!$. It is well known that

$$(1) \quad (-1)^{n-1} B_{2n} > 0 \quad (n = 1, 2, 3, \dots).$$

For proofs see for example [2, Ch. 2] and [3, Ch. 9]; another simple proof has recently been given by Mordell [1].

Moreover, if we put

$$\tan x = \sum_{n=0}^{\infty} T_{2n+1} x^{2n+1}/(2n+1)!,$$

$$\sec x = \sum_{n=0}^{\infty} (-1)^n E_{2n} x^{2n}/(2n)!,$$

it is known [2, Ch. 2] that

$$(2) \quad T_{2n+1} > 0, \quad (-1)^n E_{2n} > 0.$$

It may be of interest to note that (1) and (2) can be proved very simply in the following way. Differentiation of $\tan(\arctan x) = x$ gives

$$\tan'(\arctan x) = 1 + x^2.$$

A second differentiation gives

$$\tan''(\arctan x) = 2x(1 + x^2),$$

while a third yields

$$\tan'''(\arctan x) = (2 + 6x^2)(1 + x^2)$$

and so on. After k steps we get

$$T_k = ((1 + x^2)D)^n x \Big|_{x=0},$$

which evidently implies $T_{2n+1} > 0$.

Similarly, differentiation of $\sec(\arctan x) = \sqrt{1 + x^2}$ gives

$$\sec'(\arctan x) = x\sqrt{1 + x^2},$$

$$\sec''(\arctan x) = (1 + 2x^2)\sqrt{1 + x^2},$$

$$\sec'''(\arctan x) = (5x + 6x^3)\sqrt{1 + x^2},$$

and so on. This yields

$$E_{2n} = ((1 + x^2)D)^{2n} \sqrt{1 + x^2} \Big|_{x=0} > 0.$$

To prove (1) we note that $x \cot x - 2x \cot 2x = x \tan x$, so that

$$(3) \quad (-1)^n (1 - 2^{2n}) B_{2n} = 2n T_{2n-1} \quad (n > 0).$$

Hence (1) is implied by $T_{2n-1} > 0$.

We remark that by the method used in proving (2) we can show for example that the coefficients of $\sec^\lambda x$ are positive for $\lambda > 0$. The coefficients, except for sign, are Euler numbers of higher order [2, Ch. 6].

References

1. L. J. Mordell, The sign of the Bernoulli numbers, this MONTHLY, (preceding pages).
2. N. E. Nörlund, Vorlesungen über Differenzenrechnung, Teubner, Leipzig and Berlin, 1924.
3. J. V. Uspensky and M. A. Heaslet, Elementary Number Theory, McGraw-Hill, New York, 1939.

BASICALLY BOUNDED SETS AND A GENERALIZED HEINE-BOREL THEOREM

NEIL HINDMAN, California State College, Los Angeles

1. Introduction. The concept of boundedness plays an important role in the theory of metric spaces and has been defined in the contexts of topological vector

spaces and uniform spaces. There is thus some inherent appeal in generalizing the notion to arbitrary topological spaces. This has been explored in a general setting by Hu [1].

The advantage of the current notion is that it is defined for arbitrary topological spaces, yields a universal form of the Heine-Borel theorem, and generalizes the usual notions of boundedness for as wide a class of spaces as is possible in view of this theorem. The concept derives from an idea of Raymond Killgrove, to whom this author is indebted.

DEFINITION. A subset A of a topological space X is **basically bounded** if each basis for X has a finite subfamily covering A .

2. Generalizations of the Heine-Borel and Bolzano-Weierstrass theorems.

THEOREM 1 (Generalized Heine-Borel). *Each closed and basically bounded subset of a topological space X is compact.*

Proof. Let A be closed and basically bounded in X and let γ be an open cover of A . Let $\beta = \{U : U \text{ is open in } X \text{ and either } U \cap A = \emptyset \text{ or } U \subseteq V \text{ for some } V \text{ in } \gamma\}$. Now β is a base for X since A is closed. The finite subfamily of β which covers A guarantees that a finite subfamily of γ covers A .

THEOREM 2 (Generalized Bolzano-Weierstrass). *Each infinite basically bounded subset of a topological space has an accumulation point.*

Proof. Any infinite basically bounded set without accumulation points would be closed, hence compact, by Theorem 1.

THEOREM 3. *Each basically bounded sequence in a topological space clusters. If a sequence converges then it is basically bounded.*

3. Generalization of common boundedness notions. Recall that a subset A of a topological vector space L over a field K is said to be **bounded** if for each neighborhood V of 0 there is an element λ of K such that $A \subseteq \lambda V$. A is said to be **totally bounded** if for each neighborhood V of 0 there is a finite subset F of L such that $A \subseteq F + V$. Recall also that a subset of a uniform space X is said to be totally bounded if for each member V of the uniformity there is a finite subset F of X such that $A \subseteq V[F]$.

THEOREM 4. *Let X be a pseudo-metric space or a topological vector space. The following statements are equivalent:*

- (a) *If a subset of X is bounded (respectively totally bounded) then it is basically bounded.*
- (b) *A subset of X is bounded (respectively totally bounded) if and only if it is basically bounded.*

(c) *Each closed and bounded (respectively closed and totally bounded) subset of X is compact.*

Proof. The proof will be done in case X is a topological vector space (over the field K). The other case is similar.

(a) \rightarrow (b). Let A be a basically bounded subset of X and let V be a neighborhood of 0. Let $\beta = \{x + U : x \in X \text{ and } U \text{ is an open neighborhood of } 0 \text{ contained in } V\}$. Then β is a base for X and so A is contained in finitely many translates of V . A is thus totally bounded and hence bounded.

(b) \rightarrow (c). Any closed and bounded (respectively closed and totally bounded) subset of X is closed and basically bounded, hence compact.

(c) \rightarrow (a). Let A be a bounded (respectively totally bounded) subset of X . Then $\text{cl } A$ is bounded (respectively totally bounded) so is compact. Any compact set is basically bounded.

THEOREM 5. *Let X be a uniform space. The following statements are equivalent:*

- (a) *If a subset of X is totally bounded then it is basically bounded.*
- (b) *A subset of X is totally bounded if and only if it is basically bounded.*
- (c) *If a subset of X is closed and totally bounded then it is compact.*

Proof. (a) \rightarrow (b). Let A be a basically bounded subset of X and let U be a member of the uniformity. Let $\beta = \{W : W \text{ is open and } W \subseteq U(x) \text{ for some } x \text{ in } X\}$. Then β is a basis for X so there is a finite subfamily of β containing A . Hence $A \subseteq \bigcup_{i=1}^n U(x_i)$ for some $\{x_i\}_{i=1}^n \subseteq X$.

The rest of the proof is identical to that of Theorem 4.

As previously remarked, Theorems 4 and 5 show that the basically bounded concept generalizes these five concepts for precisely as wide a class of spaces as is possible in view of Theorem 1.

Hu [1] introduced the concept of compact bounded sets whereby a set is compact bounded if its closure is compact. Theorems 1, 2 and 3 also hold if "basically bounded" is replaced by "compact bounded". Clearly each compact bounded set is basically bounded and it is easily proved that in a regular space the converse holds. The following example shows that the basically bounded concept is indeed more general than that of compact boundedness.

Example. A Hausdorff space with a basically bounded subset which is not compact bounded.

Let X be the closed interval $[0,1]$ and let $A = X \setminus \mathcal{Q}$. Let $\Gamma = \{U \cup (V \cap A) : U \text{ and } V \text{ are open in the usual topology on } [0,1]\}$. Γ is easily checked to be a Hausdorff topology on X . To see that A is basically bounded let β be any base for X . We may write $\beta = \{U_\delta \cup (V_\delta \cap A) : \delta \in \Delta\}$. Then $\{U_\delta : \delta \in \Delta\} \cup \{V_\delta : \delta \in \Delta\}$ is an open cover of $[0,1]$ in the usual topology hence has a finite subcover $\{U_\delta : \delta \in \Delta'\} \cup \{V_\delta : \delta \in \Delta'\}$. But then $A \subseteq \bigcup_{\delta \in \Delta'} (U_\delta \cup (V_\delta \cap A))$.

Now let $X_0 \in A$ and let $\{S_n\}_{n \in \mathbb{N}}$ be a sequence of rationals converging (in the usual topology) to X_0 . Then A is a neighborhood of X_0 missing $\{S_n\}_{n \in \mathbb{N}}$ so X is not compact while $X = \text{cl } A$.

Reference

1. S.-T. Hu, Boundedness in a topological space, J. Math. Pures. Appl., 28 (1949) 287-320.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

A PROBLEM ON RATIONAL FUNCTIONS

A. K. PIZER, University of California, Los Angeles

E. Straus and W. Adams [1] have shown that a nonconstant polynomial with complex coefficients is determined by the preimages of two points. More precisely, after normalizing, we have

PROPOSITION 1 (Straus and Adams). *Let $p(z)$ and $q(z)$ be polynomials, not both constant, with coefficients in \mathbb{C} , the field of complex numbers. Assume $p(z)$ and $q(z)$ have the same set of zeros and the same set of preimages of 1, i.e.,*

$$\begin{aligned} p(z_0) = 0 &\Leftrightarrow q(z_0) = 0 \\ p(z_0) = 1 &\Leftrightarrow q(z_0) = 1 \quad \text{for } z_0 \in \mathbb{C}. \end{aligned}$$

Then $p(z) = q(z)$.

Proof. Assume $n = \deg p(z) \geq \deg q(z)$. Consider the polynomial $F(z) = p'(z)(p(z) - q(z))$. Then $\deg F(z) \leq 2n - 1$. But $F(z)$ has $2n$ zeros (counted with multiplicity) occurring at those values z_0 where $p(z_0) = 0$ or $p(z_0) = 1$. Since $p'(z) \neq 0$, we see $p(z) = q(z)$.

It is natural to ask (and in fact was asked by Straus and, independently, by J. G. Clunie) if the analogous statement is true for rational functions on the complex Riemann sphere. The normalized question reads

Question I. Let $f(z)$ and $g(z)$ be two nonconstant rational functions on the Riemann sphere having the same sets of zeros, poles, and preimages of 1. Then does it follow that $f(z) = g(z)$?

The answer to Question I is negative. In fact

$$f_0(z) = \frac{-4z^3}{(z-1)^3(z+1)}$$

and

$$g_0(z) = \frac{-4z}{(z-1)(z+1)^3}$$

satisfy the hypothesis of the question, but are not identical.

REMARK. Nonconstant rational functions are determined by the preimage of four points. The proof is analogous to that of Proposition 1. See Theorem 3 in [1].

Let us say that a pair of rational functions are *associated* if they satisfy the hypothesis of Question I. Notice that if $f(z)$ and $g(z)$ are associated and $q(z)$ is any rational function, then $f(q(z))$ and $g(q(z))$ are also associated.

Letting $\beta(z) = -1/z$, we see that $g_0(z) = f_0(\beta(z))$ and letting $\tau(z) = (iz + i)/(-z + 1)$, we find

$$f_1(z) = f_0(\tau(z)) = \frac{(z+1)^3(z-1)}{(z+i)^3(z-i)}$$

and

$$g_1(z) = g_0(\tau(z)) = \frac{(z+1)(z-1)^3}{(z+i)(z-i)^3}$$

are associated and $g_1(z) = f_1(-z)$. Thus $g_1(z) = f_1(\alpha(z))$, where α is a rotation of order 2. Letting $g_n(z) = g_1(z^n)$, $f_n(z) = f_1(z^n)$ we get an associated distinct pair $f_n(z)$, $g_n(z)$ such that $g_n(z) = f_n(\gamma(z))$, where γ is a rotation of order $2n$. Thus we see that for any fractional linear transformation μ of finite even order, there exists an associated distinct pair $f_\mu(z)$, $g_\mu(z)$ of rational functions such that $g_\mu(z) = f_\mu(\mu(z))$. This gives rise to

Question II. For every linear fractional transformation τ of finite order (it suffices to consider rotations of prime order) does there exist an associated distinct pair $f(z)$, $g(z)$ such that $g(z) = f(\tau(z))$?

A more intriguing question is

Question III. Does there exist an associated distinct pair $f(z)$, $g(z)$ which is not of the type mentioned above, i.e., for which there do not exist rational functions $F(z)$, $G(z)$, $q(z)$ and a fractional linear transformation β such that $f(z) = F(q(z))$, $g(z) = G(q(z))$ and $G(z) = F(\beta(z))$?

This work was supported in part by NSF Grant GP-28696.

Reference

1. W. Adams and E. Straus, Non-Archimedean analytic functions taking the same values at the same points, *Ill. J. Math.*, 15 (1971) 418-424.

CLASSROOM NOTES

EDITED BY ROBERT GILMER

Material for this Department should be sent to David Roselle, Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803.

A CONDITION UNDER WHICH A MAPPING IS A HOMEOMORPHISM

W. R. DERRICK, Arizona State University

The purpose of this note is to present an elementary proof, suitable for a first course in topology, of a special case of a theorem of Whyburn [1, p. 116]. The proof provides a useful application of the Jordan Curve Theorem and can be used as a first step leading into the study of light open mappings.

Consider sets in the Euclidean plane; let D denote the closed unit disk. For any function f defined on D designate points in the domain by the letter z and in the range by w . Let $\text{Bdry } D$ and $\text{Int } D$ be the boundary and interior of D respectively.

THEOREM. *Let $f:D \rightarrow D$ be a continuous function which maps $\text{Bdry } D$ homeomorphically onto $\text{Bdry } D$ and is a local homeomorphism on $\text{Int } D$. Then f is a homeomorphism.*

Proof. Since D is compact and the Euclidean plane is a Hausdorff space, we need only prove that f is one-to-one and onto.

Suppose f is not onto. Then there is a projection $\Pi:f(D) \rightarrow \text{Bdry } D$, and the mapping $f^{-1}\Pi f$ is a retraction of D onto $\text{Bdry } D$, which is impossible.

The preimage of every arc in $f(\text{Int } D)$ consists of a set of disjoint arcs in $\text{Int } D$, since otherwise we would not have a local homeomorphism at any point of intersection. Furthermore, since $f(\text{Int } D) = \text{Int } D$, the preimage of an arc meeting $\text{Bdry } D$ only at an endpoint w^* consists of arcs meeting only at $f^{-1}(w^*)$.

Suppose z_0, z'_0 are distinct preimages of w_0 and a is the radial arc joining w_0 to w^* , its nearest point on $\text{Bdry } D$. (For $w_0 = 0$ take any radial arc.) By the continuity of f there exist arcs a_0, a'_0 joining z_0, z'_0 to z^* , the preimage of w^* , respectively, such that $f(a_0) = f(a'_0) = a$. Select w_1, w_2 on $\text{Bdry } D$ and denote by z_1, z_2 their preimages, let b and c be the straight line arcs joining w_0 to w_1 and w_2 , respectively, and designate by b_1, b'_1 the arcs joining z_0, z'_0 to z_1 which satisfy $f(b_1) = f(b'_1) = b$, and by c_2, c'_2 the arcs joining z_0, z'_0 to z_2 such that $f(c_2) = f(c'_2) = c$. Let d_1 and d_2 be the arcs on $\text{Bdry } D$ with endpoints z_1, z^* and z_2, z^* and satisfying $d_1 \cap d_2 = z^*$. Then by the Jordan Curve Theorem z'_0 lies inside the simple closed curve $d_1 \cup b_1 \cup c_2 \cup d_2$ because otherwise a'_0 would meet b_1 or c_2 . If z'_0 lies inside the simple closed curve $a_0 \cup d_1 \cup b_1$, then c'_2 meets a_0 or b_1 , and if z'_0 lies inside $a_0 \cup c_2 \cup d_2$, then

b_1 meets a_0 or c_2 . In either case we have a contradiction, thus each point in $f(D)$ has a unique preimage.

Reference

1. G. T. Whyburn, An open mapping approach to Hurwitz's theorem, *Trans. Amer. Math. Soc.*, 71 (1951) 113–119.

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to Shirley Hill, Department of Mathematics, University of Missouri, Kansas City, MO 64110, or to Paul Mielke, Department of Mathematics, Wabash College, Crawfordsville, IN 47933.

INDEPENDENT STUDY FOR UNDERGRADUATES

W. C. RAMALEY, Colorado College

In college catalogs there often occurs a listing for the Mathematics Department of "Independent Study". A great many diverse activities occur under this rubric. What follows is about some of these activities and their importance to the students, to the college, and perhaps, to the graduate school. The conclusions drawn come from my experience over the past 5 years at a college which devotes all its attention to a quality undergraduate education.

At Carleton College "Independent Study" has had a long and dynamic history [4, 5]. In the 1970–71 academic year 23 students enrolled in the course. Their work could be classified as follows: 5 read a regularly offered course in a term the course was not offered, 4 read from special bibliographies prepared for a "reading course" (usually in history), 5 covered material that would be treated in an advanced course if Carleton could offer that course, 5 did directed research which may or may not have been related to a previous course, and 4 did truly independent research. Over the past 5 years the proportions and total numbers have remained fairly constant.

Each type of activity makes different demands on the student and on the professor who supervises the activity. For a regular course being covered in a term when it is not taught, a faculty member who has taught the course may be able to spend only an hour or so a week with the student in addition to preparing and grading a final examination. A faculty member who has not taught the course should avoid this sort of independent study, unless he wants to learn the material himself. Even then, "learning-by-teaching" has obvious limitations that must be carefully considered.

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

ASSOCIATE EDITORS: JOSHUA BARLAZ, ERIC S. LANGFORD. COLLABORATING EDITORS: LEONARD CARLITZ, GULBANK D. CHAKERIAN, HASKELL COHEN, S. ASHBY FOOTE, ISRAEL N. HERSTEIN, MURRAY S. KLAMKIN, DANIEL J. KLEITMAN, ROGER C. LYNDON, MARVIN MARCUS, CHRISTOPH NEUGEBAUER, ALBERT WILANSKY, AND UNIVERSITY OF MAINE PROBLEMS GROUP: EARL M. L. BEARD, GEORGE S. CUNNINGHAM, CLAYTON W. DODGE, OSKAR FEICHTINGER, WILLIAM R. GEIGER, RAMESH GUPTA, GARY HAGGARD, PHILIP M. LOCKE, JOHN C. MAIRHUBER, CURTIS S. MORSE, GRATTAN P. MURPHY, EDWARD S. NORTHAM AND WILLIAM L. SOULE, JR.

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of elementary Problems in this issue should be typed (with double spacing) and should be mailed before August 31, 1973.

E 2414. *Proposed by J. G. Wendel, University of Michigan*

In one form of chess match $2n$ games are played, wins count 1 point each, draws $\frac{1}{2}$, losses are worth 0. In order to win the match, the defender needs only score at least n , while the challenger must achieve at least $n + \frac{1}{2}$. Suppose that the two players are of equal strength, and that the probability of a draw is a constant δ . Prove or disprove: the defender's chance of keeping his title is an increasing function of δ .

E 2415. *Proposed by C. D. H. Cooper, Macquarie University, Australia*

Find all positive integers n having the property that each positive divisor (> 1) of n has the form $a^r + 1$ where a, r are integers and $r > 1$.

E 2416. *Proposed by F. T. Howard, Wake Forest University*

Let p_1, \dots, p_k be distinct primes, e_1, \dots, e_k arbitrary non-negative integers, and r a fixed positive integer. Prove that there are infinitely many positive integers n with the property that $p_i^{e_i}$ ($i = 1, 2, \dots, k$) is the highest power of p_i which divides the binomial coefficient $\binom{n}{r}$.

E 2417. *Proposed by Ioan Tomescu, University of Bucharest, Rumania*

The number of ways of filling a $2 \times n$ rectangle with dominoes (i.e. with 1×2

rectangles) is well known (see problem E 1470 [1962, 61]). On page 139 of his book *Polyominoes*, S. W. Golomb asks for the corresponding result for $3 \times n$ rectangles.

Let $\mathcal{U}(n)$ be the number of ways of covering a $3 \times n$ rectangle with dominoes. Obviously $\mathcal{U}(n) = 0$ if n is odd. Show that

$$\mathcal{U}(2m) = \frac{1}{2\sqrt{3}} [(\sqrt{3} + 1)(2 + \sqrt{3})^m + (\sqrt{3} - 1)(2 - \sqrt{3})^m].$$

E 2418. *Proposed by C. A. Nicol, University of South Carolina*

Characterize those subsets S of the natural numbers with the property that every sum of elements taken from S (repetitions allowed) is composite.

E 2419. *Proposed by A. W. Walker, Toronto, Canada*

Points G, H, I, O are the centroid, orthocenter, incenter and circumcenter of a scalene triangle Δ , N and P are the midpoints of line segments OH and IH , F is the contact point of the incircle and nine-point circle of Δ , E is the reflection of F in the right bisector of OH , and L is the inverse of E in the circle on GH as diameter. Prove:

- (a) F and I are inverse in the circle with center N , radius NP ;
- (b) if $OI = \sqrt{3} \cdot OG$, points L and I coincide;
- (c) lines FG, IL, OP concur.

SOLUTIONS OF ELEMENTARY PROBLEMS

A Test for Primality

E 2355 [1972, 518]. *Proposed by Arthur Marshall, Madison, Wisconsin*

Given any odd integer $n > 3$, let k and j be the smallest natural numbers such that $kn + 1$ and jn are squares. Prove that n is prime if and only if both k and j are greater than $n/4$.

Solution by R. J. Evans, Jackson State College, Mississippi. All variables represent natural numbers. Suppose n is prime. Then $n \nmid j$ so that $j \geq n > n/4$. Also $kn = (a - 1)(a + 1)$ for some a , so that for some b , $a \pm 1 = nb$. Thus $kn \geq nb(nb - 2) \geq n(n - 2)$ so $k \geq n - 2 > n/4$. Conversely, suppose n is composite. If n is a prime power, $n = p^r$, then $j = 1$ or $j = p$ according as r is even or odd. In either case, $j < n/4$, the desired result. Now suppose n is not a prime power, so there exist relatively prime odd integers p and q greater than 1 such that $n = pq$. By the Chinese Remainder Theorem there exists an x such that $1 < x < n - 1$, $p \mid (x - 1)$ and $q \mid (x + 1)$. Let $a = \min(x, n - x)$. Then $a^2 - 1 = k_1 n$ for some k_1 . Thus $k_1 n < a^2 < (n/2)^2$, so $k \leq k_1 < n/4$.

Also solved by Problem Solving Group, Berne (Switzerland), John Christopher, A. P. Geist, M. G. Greening (Australia), C. V. Heuer & G. A. Heuer, Wells Johnson, L. Kuipers, O. P. Lossers

(Netherlands), Carolyn MacDonald (partial solution), Helen M. Marston, L. E. Mattics, M. R. Modak (India), Kenneth Schilling, Nan-Shan Shou, E. P. Starke, Allen Stenger, Charles Wexler, and the proposer.

Editor's Comment. Shou generalized the problem and its solution by showing that the conclusion is valid for 2 and for any integer of the form p or $2p$, p an odd prime. Charles Wexler noted that since necessary and sufficient conditions for primality other than the definition and Wilson's theorem are very rare indeed, the problem is of more than passing interest.

The Cancellation Law for Convex Sets

E 2358 [1972, 519]. *Proposed by W. H. Ruckle, Clemson University*

Suppose that A and B are closed, convex sets and that C is bounded. Show that if $A + C = B + C$, then necessarily $A = B$.

Editor's Comment. Both Richard Laatch and George Painter note that this problem is Lemma 2, p. 167 of Hans Radström, *An embedding theorem for spaces of convex sets*, Proc. Amer. Math. Soc., 3 (1952) 165–169. It should be noted that the elementary proof given in this reference, holds without change in a (real or complex) linear topological space. (See H. H. Schaefer, *Topological Vector Spaces*, Macmillan, New York, 1964, 25–27.)

Many solvers observe that obviously C must be nonempty.

P. J. Zwier and the proposer provide examples which show that the assumptions on A , B and C are necessary. If C were not bounded, we could take C to be the real line with $A = \{0\}$ and $B = \{1\}$. If A were not convex, we could take $A = \{0, 1\}$ and $B = C = [0, 1]$. If A were not closed, we could take $A = C = (0, 1)$ and $B = [0, 1]$.

Also solved by Sheldon Axler, Ronald Evans, C. V. Heuer & G. A. Heuer, E. M. Klein, O. P. Lossers (Netherlands), Simeon Reich (Israel), Peter Renz, Ralph Seifert, Walter Stromquist, and Jack Zilver.

Circular Regions Determined by Chords

E 2359 [1972, 519]. *Proposed by T. C. Brown, Simon Fraser University, Burnaby, Canada*

Place n distinct points on the circumference of a circle and draw all possible chords through pairs of these points. Assume no three chords are concurrent and let a_n denote the resulting number of regions within the circle. Then the sequence a_1, a_2, \dots begins 1, 2, 4, 8, 16, 31, \dots . What is a_n in general?

Solution by Norman Bauman, Nanuet, N. Y. Consider the more general problem of a region crossed by l lines with p interior points of intersection. One easily shows by induction that the number of disjoint subregions created is $p + l + 1$. In the special case of the problem, n points about a circle determine $\binom{n}{2}$ lines and $\binom{n}{4}$ internal intersections. The answer is, therefore, $a_n = \binom{n}{2} + \binom{n}{4} + 1$.

Also solved by 54 other readers.

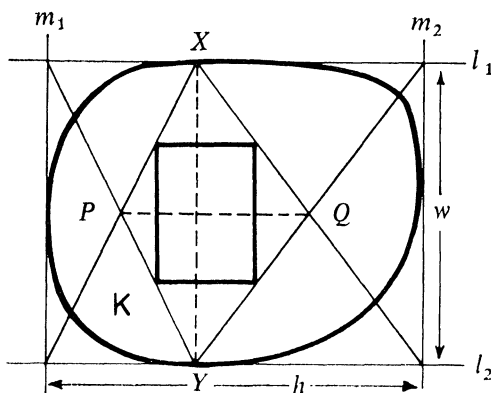
Editor's Note. The problem is certainly not new. Readers point out that it has appeared in at least ten journals and books as well as the 1967 Santa Clara mathematics contest for high school students. The problem appears in Yaglom and Yaglom, *Challenging Mathematical Problems*, Holden-Day, 1964, p. 108; in T. Murphy, *The dissection of a circle by chords*, *The Mathematical Gazette*, #396 (1972), pp. 113–115; and in Jay Graening, *Induction, fallible but valuable*, *The Mathematics Teacher*, Feb. 1971, pp. 127–131.

Subrectangles of a Convex Body

E 2360 [1972, 519]. *Proposed by G. D. Chakerian, University of California, Davis*

A convex body in the plane is a convex set with non-empty interior. The width of a convex body is the minimum possible distance between parallel supporting lines. Show that if K is a convex body in the plane of width w and area A , then K contains a rectangle with dimensions $\sqrt{A}/4$ by $w/2$.

I. Solution by G. A. Converse and J. E. Wetzel, University of Illinois. Let l_1 and l_2 be two parallel support lines to K at minimum distance w apart, and suppose that the two perpendicular support lines m_1 and m_2 are h apart. Since $A \leq hw \leq h^2$, evidently $h \geq \sqrt{A}$. There are contact points X and Y of K on l_1 and l_2 so that $XY \perp l_1$ (by an argument similar to that given on p. 117 of Yaglom and Boltyanskiĭ, *Convex Figures*, Holt, Rinehart and Winston, 1961). Let P and Q be the intersections of the diagonals of the two rectangles with edge XY (see the figure). Then the quadrilateral $XPYQ$ lies in K , $PQ \perp XY$, $XY = w$, and $PQ = h/2$. The rectangle whose vertices are the midpoints of the sides of the quadrilateral $XPYQ$ lies in K and has sides $w/2$ and $h/4$; it is worth noting that it has area $T = wh/8 \geq A/8$. Thus K surely contains the smaller rectangle with sides $w/2$ and $\sqrt{A}/4$; and, moreover, the side of length $w/2$ can be chosen to lie perpendicular to the support line l_1 .



II. Solution by the proposer. We establish a sharper result, namely, K contains a rectangle with dimensions $\sqrt{A}/2$ by $w/2$.

DEFINITION. A *diameter* of K is any chord whose endpoints lie on two parallel supporting lines.

LEMMA. There exists a quadrilateral $PQRS$ inscribed in K such that PR and QS are diameters of K , with PR orthogonal to QS , and such that supporting lines of K through the vertices of the quadrilateral form a rectangle circumscribed about K (the sides of the rectangle need not be parallel to the diameters).

Proof. We prove the lemma in case K is smooth and strictly convex. The general result follows by standard approximation arguments. In this case, to each diameter there corresponds a unique orthogonal diameter and a unique circumscribed parallelogram whose sides contain the endpoints of the diameters. If for some choice of direction for the initial diameter the corresponding parallelogram is not a rectangle, it is evident that a "rotation" of this configuration through 90° , interchanging the roles of the two diameters, will by continuity yield our circumscribed rectangle in some intermediate position.

The midpoints of the sides of $PQRS$ are the vertices of a rectangle J contained in K . If $a \geq b$ are the lengths of the sides of the rectangle circumscribed about K (with sides passing through P, Q, R, S), then the sides of J have length at least $a/2$ and $b/2$ respectively. Since $a \geq b \geq w$, we see that both sides of J have length at least $w/2$. Also, $A \leq ab \leq a^2$, so $a/2 \geq \sqrt{A}/2$. Thus K contains a rectangle of the required dimensions.

Partition Permutations

E 2364 [1972, 663]. *Proposed by G. J. Michaelides, University of South Florida*

Suppose that r is a positive integer and that (i_1, i_2, \dots, i_n) is a partition of r into nonnegative integers. Show that if p is a prime factor of n which is relatively prime to r , then the number of (distinct) permutations of (i_1, i_2, \dots, i_n) is divisible by p .

Solution by D. M. Bloom, Brooklyn College. Let the partition consist of k distinct values j_1, j_2, \dots, j_k having respective multiplicities m_1, m_2, \dots, m_k . (Thus $m_1 j_1 + m_2 j_2 + \dots + m_k j_k = r$ and $m_1 + m_2 + \dots + m_k = n$.) Since $p \nmid r$, at least one of the m 's (say m_1) also is not divisible by p . Since $p \mid n$, it follows that $p \mid \binom{n}{m_1}$. But the number N of distinct partitions is given by

$$N = \binom{n}{m_1} \binom{n - m_1}{m_2} \dots \binom{n - m_1 - \dots - m_{k-1}}{m_k}.$$

Thus $p \mid N$.

Also solved by Irl Bivens, John Christopher, Ellen Hertz, Joseph Hoffman, F. T. Howard, Wells Johnson, James King, Harry Lass, M. R. Modak (India), Paul Stockmeyer, and the proposer.

Editor's Comment. Stockmeyer dispenses with the requirement that p be prime by arguing that each prime-power factor q^a of p must divide N . Howard establishes a similar result.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers — The State University, New Brunswick, N. J., 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before August 31, 1973. Contributors (in the United States) who desire acknowledgement of receipt of their solutions are asked to enclose self-addressed, stamped postcards.

An asterisk () means neither the proposer nor the editors supplied a solution.*

5911. *Proposed by Bill Knight, California Institute of Technology*

Let F_n be the n th term of the sequence defined by

$$F_n = (n + 2)F_{n-1} - (n - 1)F_{n-2}, \quad F_1 = a, \quad F_2 = b.$$

Find an explicit formula for F_n .

5912. *Proposed by R. B. Kirk, Southern Illinois University*

Let X be a compact Hausdorff space, and let C denote the space of continuous functions on X . Assume that C can be written as a countable union of equicontinuous sets. Prove that X is finite.

5913. *Proposed by D. E. Daykin and J. K. Dugdale, University of Reading, England*

Let H be a real or complex Hilbert space and let x, y, z be points in H . We call $\langle x, y, z \rangle$ a triangle. As usual

$$L(y, z) = \{y + \alpha(y - z) : \alpha \text{ a scalar}\}$$

is the line through y and z : and the distance of x from $L(y, z)$ is

$$p(x, L(y, z)) = \inf \{\|x - w\| : w \in L(y, z)\}.$$

For convenience put $a = \|x - y\|$, $b = \|y - z\|$, $c = \|z - x\|$ and $s = \frac{1}{2}(a + b + c)$.

Euclidean geometry suggests two definitions

$$A_1 = \frac{1}{2}p(x, L(y, z)) \|y - z\|$$

and

$$A_2 = \sqrt{s(s - a)(s - b)(s - c)}$$

for the area of triangle $\langle x, y, z \rangle$. Compare the values of A_1 and A_2 and determine when they are equal.

5914. *Proposed by D. E. Daykin, C. E. Linderholm and Albert Wilansky, University of Reading, England.*

Show that if $A = \{z_1, z_2, \dots, z_n\}$ is a finite set of n complex numbers, there is a subset B of A such that

$$\left| \sum_{z \in B} z \right| \geq \pi^{-1} \sum_{1 \leq i \leq n} |z_i|.$$

5915*. *Proposed by D. M. Battany, Oceanside, California*

Let p_n be the n th prime. Show that

$$\prod_{p < p_1 \cdots p_n} \left(1 - \frac{1}{p} \right) \leq \frac{1}{p_n}$$

for all prime p , or isolate the exceptional values.

SOLUTIONS OF ADVANCED PROBLEMS

Discontinuity in a Function with All Partial Derivatives

5840 [1972, 187]. *Proposed by Maury Horowitz, Nick Metas and Gerald Leibowitz, University of Connecticut*

Can one construct a real-valued function f whose domain is an open set U in R^2 such that f has all partial derivatives of all orders at every point in U yet there is some point in U at which f is not continuous?

Solution by Wolfe Snow, Brooklyn College. Let

$$f(x, y) = \begin{cases} \frac{\exp(x^{-2}y^{-2})}{\exp(x^{-4}) + \exp(y^{-4})} & \text{for } xy \neq 0, \\ 0 & \text{for } xy = 0. \end{cases}$$

Then f is not continuous at $(0, 0)$ since $\lim_{x=y \rightarrow 0} f(x, y) = \frac{1}{2}$.

The only possible source of differentiation difficulty is when x or y is 0. This does not pose any problem, however, since for all derivatives the exponential in the denominator dominates the exponential in the numerator and any power terms that may arise, and consequently all derivatives are 0 when either x or y is 0.

Thus, f has all partial derivatives of all orders at every point of R^2 , yet f is not continuous at $(0, 0)$.

Also solved by R. T. Baumel, J. M. Bell, R. L. Bishop, A. A. Blank, R. A. Christiansen, L. E. Clarke (England), S. Cullinane, J. Diederich, A. G. Dors, G. J. Foschini, G. Freilich, R. Katz, I. Korec (Yugoslavia), H. C. Kranzer, O. P. Lossers (Netherlands), M. Machover, L. Mattics, J. G. Mauldon, P. L. Montgomery, C. J. Neugebauer, S. Rajnak, J. Rätz, (Switzerland), H. Van Evelghem (Belgium), A. Weinmann (England), and A. C. Williams.

Notes. Bishop notes an example on p. 20 of Bishop and Goldberg, *Tensor Analysis on Manifolds*. Dors cites a method of construction which is found in Appendix L of E. E. Moise, *Calculus*. Blank and V. Mizel note an example of a function with a discontinuity, yet having directional derivatives of all orders.

Fourier-Stieltjes Transform of a Continuous Measure

5841 [1972, 187]. *Proposed by L.-S. Hahn, University of New Mexico*

Is there a (complex) continuous measure (i.e., $\mu(E) = 0$ if E is countable) on the real line, whose Fourier-Stieltjes transform has modulus 1 everywhere on the real line?

Solution by G. M. Leibowitz, University of Connecticut. There is no such measure. We offer two proofs, one quantitative, the other qualitative.

I. By a theorem of Wiener, if $\mu \in M(\mathbb{R})$, then

$$\sum |\mu(\{x\})|^2 = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T |\hat{\mu}(t)|^2 dt.$$

(See Katznelson, *An Introduction to Harmonic Analysis*, p. 138.) So if $\hat{\mu}$ has modulus 1, $\sum |\mu(\{x\})|^2 = 1$, and μ is not continuous. (An analogous averaging process holds on any locally compact abelian group, yielding the same result.)

II. Let $\mu \in M(G)$, G any LCA group. Set $\tilde{\mu}(E) = \mu(-\bar{E})$. Assume that $|\hat{\mu}| = 1$. Then $(\mu * \tilde{\mu})^2 = |\hat{\mu}|^2 = 1 = \hat{\delta}$, where δ is the point mass at 0. Hence $\mu * \tilde{\mu} = \delta$ by uniqueness of Fourier-Stieltjes transforms. The continuous measures form an ideal, so μ is not continuous.

We could also quote Corollary 5.6.9(b) in Rudin, *Fourier Analysis on Groups*.

Also solved by R. W. Chaney, S. H. Friedberg, C. C. Graham, D. Lind, N. X. Uy, and the proposer.

Orthogonal Projections

5842 [1972, 307]. *Proposed by B. B. Winter, Eugene, Oregon*

Let T be a linear (not necessarily continuous) map of a Hilbert space H to itself. Suppose there exists a subset S such that $Tx \in S$ and $x - Tx \in S^\perp$ for all $x \in H$. Show that S is a closed linear subspace and that T is the (necessarily continuous) orthogonal projection of H onto S .

Solution by S. P. Gudder, University of Denver. If $x \in S$, since $Tx \in S$ and $x - Tx \in S^\perp$, we have $(x - Tx) \perp (x - Tx)$, so $Tx = x$. We thus see that $x \in S$ if and only if $Tx = x$. Hence $T^2y = T(Ty) = Ty$ for all $y \in H$ and T is a projection. Now T is hermitian since

$$(\acute{x}, Ty) = ((x - Tx) + Tx, Ty) = (Tx, Ty) = (Tx, (Ty - y) + y) = (Tx, y)$$

for all $x, y \in H$. Since a hermitian operator defined on all the space is self-adjoint, T is a self-adjoint projection and hence an orthogonal projection. Since S is the range of T , S is a closed subspace.

Also solved by forty-two other contributors.

A First Order Non-linear Differential Inequality

5843 [1972, 307]. *Proposed by N. P. Callas, Office of Scientific Research, U.S. Air Force*

Show that if $\sigma(x) \geq 0$ satisfies the nonlinear differential inequality

$$\sigma'(x) + b(x)\sigma(x) \leq f(x)[\sigma(x)]^\alpha,$$

where $\sigma(a) = c$ and $0 \leq \alpha < 1$, then

$$\sigma(x) \leq \exp\left(-\int_a^x b(\tau)d\tau\right) \left[\int_a^x (1-\alpha)f(\tau) \exp\left(\int_a^\tau (1-\alpha)b(t)dt\right)d\tau + c^{1-\alpha}\right]^{1/(1-\alpha)}$$

Solution by D. G. Belanger, University of South Alabama. Assume that $\sigma(x) > 0$ on $[a, x]$. Divide the inequality by $[\sigma(x)]^\alpha$ and make the substitution $v(x) = (\sigma(x))^{1-\alpha}$ obtaining

$$\frac{1}{1-\alpha}v'(x) + b(x)v(x) \leq f(x).$$

We now multiply by the integrating factor $(1-\alpha)\exp(\int_a^x (1-\alpha)b(\tau)d\tau)$ obtaining

$$\frac{d(v(x) \exp(\int_a^x (1-\alpha)b(\tau)d\tau))}{dx} \leq (1-\alpha)f(x) \exp\left(\int_a^x (1-\alpha)b(\tau)d\tau\right).$$

Integrating:

$$\begin{aligned} v(x) \exp\left(\int_a^x (1-\alpha)b(\tau)d\tau\right) &\leq \int_a^x \left[(1-\alpha)f(\tau) \exp\left(\int_a^\tau (1-\alpha)b(\tau)d\tau\right)\right] dt + v(a), \\ v(x) &\leq \exp\left(-\int_a^x (1-\alpha)b(\tau)d\tau\right) \cdot \\ &\quad \left\{ \int_a^x \left[(1-\alpha)f(t) \exp\left(\int_a^t (1-\alpha)b(\tau)d\tau\right)\right] dt + v(a) \right\}. \end{aligned}$$

Since $v(x) = [\sigma(x)]^{1-\alpha}$,

$$\begin{aligned} [\sigma(x)]^{1-\alpha} &\leq \exp\left(-\int_a^x (1-\alpha)b(\tau)d\tau\right) \cdot \\ &\quad \left\{ \int_a^x \left[(1-\alpha)f(\tau) \exp\left(\int_a^\tau (1-\alpha)b(\tau)d\tau\right)\right] dt + c^{1-\alpha} \right\}, \\ \sigma(x) &\leq \exp\left(-\int_a^x b(\tau)d\tau\right) \cdot \end{aligned}$$

$$\left\{ \int_a^x \left[(1-x)f(t) \exp \left(\int_a^t (1-x)b(\tau)d\tau \right) \right] dt + c^{1-x} \right\}^{1/(1-x)}.$$

Also solved by J. E. Chance, F. A. Homann, S. J., A. A. Jagers (Netherlands), G. A. Kemper, Charlotte Krauthammer (Austria), J. R. Kuttler, Beatriz Margolis (Argentina), R. J. Schaar, J. S. Shipman, T. Teichmann, H. C. Wente, and the proposer.

The original statement contained a misprint, as discovered by all contributors: as first printed, $b(\tau)$ had an incorrect coefficient $(1-a)$.

Discontinuities of Functions in \mathbb{R}^2

5844 [1972, 307]. *Proposed by L.-S. Hahn, University of New Mexico*

Construct a function defined everywhere in the plane which is nowhere continuous and yet is continuous in each variable separately, or prove such a function does not exist.

Solution by G. M. Leibowitz, University of Connecticut. In volume one of E. W. Hobson, *The Theory of Functions of a Real Variable*, reprinted by Dover, 1957, we see on p. 449 that if f is separately continuous in each variable, then f is continuous at points on each graph of a continuous function. Hence no such function exists.

Also solved by Bruce Ferrero, O. P. Lossers (Netherlands), C. J. Neugebauer, T. Šalát (Czechoslovakia), and the proposer.

Note. We are referred by Šalát to F. W. Carroll, *Separately continuous functions*, this MONTHLY, V. 78 (1971), p. 175; and by Lossers to C. Goffman, *Real Functions*. The critical fact is that f is in the first Baire class.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN

with the assistance of the mathematics departments of St. Olaf and Carleton Colleges

COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER, CARLETON COLLEGE

Printed materials for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057. Films and correspondence relating to films should be sent to Seymour Schuster, Carleton College, Northfield MN 55057.

All unsigned material is written by the editors. A boldface capital C in the margin indicates that a review is based in part on classroom use. Professors willing to write such a review should inform the editor in order to avoid duplication.

Perspectives in Mathematics. By David E. Penney. Benjamin, Menlo Park, California, 1972. xiv + 349 pp. \$9.95. (Telegraphic Review, April 1972.)

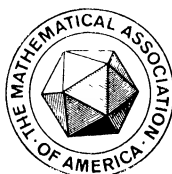
Mathematics in Civilization. By H. L. Resnikoff and R. O. Wells, Jr. Holt, Rinehart,

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF
THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 80



NUMBER 6

CODEN: AMMYAE

PART I

CONTENTS

A History of the Prime Number Theorem	L. J. GOLDSTEIN	599
Differentiation under the Integral Sign	HARLEY FLANDERS	615
The Stanford University Competitive Examination in Mathematics		
.	G. POLYA AND J. KILPATRICK	627
How to Classify Differential Polynomials	REUBEN HERSH	641
The Cesàro Operators and their Generalizations: Examples in Infinite-Dimensional Linear Analysis	GERALD LEIBOWITZ	654
A. A. Albert	D. ZELINSKY	661

MATHEMATICAL NOTES

Functions Satisfying a Mean Value Property at their Zeros	D. P. STANFORD	665
On an Extension of the Theorem of Hausdorff-Young	LIANG-SHIN HAHN	667
A Characterization of the $n \times n$ Matrices over a Finite Field		
.	J. V. BRAWLEY AND L. CARLITZ	670
Another Proof of Bernstein's Theorem	P. J. O'HARA	673
Addendum to "On the Diffeomorphisms of Euclidean Space"	W. B. GORDON	674

RESEARCH PROBLEMS

How Unexpected is the Prime Number Theorem?	M. D. HIRSCHHORN	675
-------------------------------------------------------	------------------	-----

CLASSROOM NOTES

The Indecomposability of the Dyadic Solenoid	S. B. NADLER, JR.	677
The Differentiability Properties of Typical Functions in $C[a, b]$	A. M. BRUCKNER	679
Representing a Finite Borel Measure in Terms of its Distribution Function	J. J. HIGGINS	683

(Continued on inside cover)

JUNE-JULY

1973

MATHEMATICAL EDUCATION

Using Student-tutors in Precalculus Instruction	T. A. EISENBERG AND J. B. BROWNE	685
Economics as a Minor for Undergraduate Mathematics Majors.	D. F. ELLIS	688
Survival for Mathematics Students!	B. B. HUGHES	689
ELEMENTARY PROBLEMS AND SOLUTIONS		691
ADVANCED PROBLEMS AND SOLUTIONS.		697
REVIEWS		702
NEWS AND NOTICES		721
MATHEMATICAL ASSOCIATION OF AMERICA		722
February Meeting of the Northern California Section.		722
November Meeting of the Maryland-District of Columbia-Virginia Section.		722
November Meeting of the Philadelphia Section		723
Calendars of Future Meetings		724

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss.

Backlog: Main Articles 12 months, Math. Notes 13 months, Research Problems 7 months, Classroom Notes 11 months, Math. Education 10 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to ALEX ROSENBERG, Department of Mathematics, Cornell University, Ithaca, N.Y. 14850; NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D.C. 20036.

HARLEY FLANDERS, *Editor*

ALEX ROSENBERG, *Editor-Elect*

ASSOCIATE EDITORS

JOSHUA BARLAZ	J. G. HARVEY	SEYMOUR SCHUSTER
E. R. BERLEKAMP	ERIC S. LANGFORD	J. ARTHUR SEEBACH, Jr.
JANE W. DI PAOLA	P. D. LAX	E. P. STARKE
ROBERT GILMER	ARTHUR MATTUCK	LYNN A. STEEN
RICHARD GUY	M. W. POWNALL	JAMES WENDEL
RAOUL HAILPERN	GIAN-CARLO ROTA	

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are \$12.50. For nonmembers the subscription price is \$18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June-July, August-September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices.

Copyright © The Mathematical Association of America (Incorporated), 1973

PRINTED IN THE UNITED STATES OF AMERICA

A HISTORY OF THE PRIME NUMBER THEOREM

L. J. GOLDSTEIN, University of Maryland

The sequence of prime numbers, which begins

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, \dots,$$

has held untold fascination for mathematicians, both professionals and amateurs alike. The basic theorem which we shall discuss in this lecture is known as the **prime number theorem** and allows one to predict, at least in gross terms, the way in which the primes are distributed. Let x be a positive real number, and let $\pi(x)$ = the number of primes $\leq x$. Then the prime number theorem asserts that

$$(1) \quad \lim_{x \rightarrow \infty} \frac{\pi(x)}{x/\log x} = 1,$$

where $\log x$ denotes the natural log of x . In other words, the prime number theorem asserts that

$$(2) \quad \pi(x) = \frac{x}{\log x} + o\left(\frac{x}{\log x}\right), \quad (x \rightarrow \infty),$$

where $o(x/\log x)$ stands for a function $f(x)$ with the property

$$\lim_{x \rightarrow \infty} \frac{f(x)}{x/\log x} = 0.$$

Actually, for reasons which will become clear later, it is much better to replace (1) and (2) by the following equivalent assertion:

$$(3) \quad \pi(x) = \int_2^x \frac{dy}{\log y} + o\left(\frac{x}{\log x}\right).$$

To prove that (2) and (3) are equivalent, it suffices to integrate

$$\int_2^x \frac{dy}{\log y}$$

once by parts to get

$$(4) \quad \int_2^x \frac{dy}{\log y} = \frac{x}{\log x} - \frac{2}{\log 2} + \int_2^x \frac{dy}{\log^2 y}.$$

Larry Goldstein received his Princeton Ph.D. under G. Shimura. After a Gibbs instructorship at Yale, he joined the Univ. of Maryland as Associate Professor and now is Professor. His research is in Analytic and Algebraic Number Theory and Automorphic Functions. He is the author of *Analytic Number Theory* (Prentice-Hall 1971), and *Abstract Algebra, A First Course* (Prentice-Hall, to appear). *Editor*.

However, for $x \geq 4$,

$$\begin{aligned}
 \int_2^x \frac{dy}{\log^2 y} &= \int_2^{\sqrt{x}} \frac{dy}{\log^2 y} + \int_{\sqrt{x}}^x \frac{dy}{\log^2 y} \\
 (5) \qquad &\leq \sqrt{x} \cdot \frac{1}{\log^2 2} + x \cdot \frac{1}{\log^2 (\sqrt{x})} \\
 &= o\left(\frac{x}{\log x}\right),
 \end{aligned}$$

where we have used the fact that $1/\log^2 x$ is monotone decreasing for $x > 1$. It is clear that (4) and (5) show that (2) and (3) are equivalent to one another. The advantage of the version (3) is that the function

$$\text{Li}(x) = \int_2^x \frac{dy}{\log y},$$

called the **logarithmic integral**, provides a much closer numerical approximation to $\pi(x)$ than does $x/\log x$. This is a rather deep fact and we shall return to it.

In this lecture, I should like to explore the history of the ideas which led up to the prime number theorem and to its proof, which was not supplied until some 100 years after the first conjecture was made. The history of the prime number theorem provides a beautiful example of the way in which great ideas develop and interrelate, feeding upon one another ultimately to yield a coherent theory which rather completely explains observed phenomena.

The very conception of a prime number goes back to antiquity, although it is not possible to say precisely when the concept first was clearly formulated. However, a number of elementary facts concerning the primes were known to the Greeks. Let us cite three examples, all of which appear in Euclid:

(i) (*Fundamental Theorem of Arithmetic*): Every positive integer n can be written as a product of primes. Moreover, this expression of n is unique up to a rearrangement of the factors.

(ii) There exist infinitely many primes.

(iii) The primes may be effectively listed using the so-called “sieve of Eratosthenes”.

We will not comment on (i), (iii) any further, since they are part of the curriculum of most undergraduate courses in number theory, and hence are probably familiar to most of you. However, there is a proof of (ii) which is quite different from Euclid’s well-known proof and which is very significant to the history of the prime number theorem. This proof is due to the Swiss mathematician Leonhard Euler and dates from the middle of the 18th century. It runs as follows:

Assume that p_1, \dots, p_N is a complete list of all primes, and consider the product

$$(6) \quad \prod_{i=1}^N \left(1 - \frac{1}{p_i}\right)^{-1} = \prod_{i=1}^N \left(1 + \frac{1}{p_i} + \frac{1}{p_i^2} + \cdots\right).$$

Since every positive integer n can be written uniquely as a product of prime powers, every unit fraction $1/n$ appears in the formal expansion of the product (6). For example, if $n = p_1^{a_1} \cdots p_N^{a_N}$, then $1/n$ occurs from multiplying the terms

$$1/p_1^{a_1}, 1/p_2^{a_2}, \dots, 1/p_N^{a_N}.$$

Therefore, if R is any positive integer,

$$(7) \quad \prod_{i=1}^N \left(1 - \frac{1}{p_i}\right)^{-1} \geq \sum_{n=1}^R 1/n.$$

However, as $R \rightarrow \infty$, the sum on the right hand side of (7) tends to infinity, which contradicts (7). Thus, p_1, \dots, p_N cannot be a complete list of all primes. We should make two comments about Euler's proof: First, it links the Fundamental Theorem of Arithmetic with the infinitude of primes. Second, it uses an analytic fact, namely the divergence of the harmonic series, to conclude an arithmetic result. It is this latter feature which became the cornerstone upon which much of 19th century number theory was erected.

The first published statement which came close to the prime number theorem was due to Legendre in 1798 [8]. He asserted that $\pi(x)$ is of the form $x/(A \log x + B)$ for constants A and B . On the basis of numerical work, Legendre refined his conjecture in 1808 [9] by asserting that

$$\pi(x) = \frac{x}{\log x + A(x)},$$

where $A(x)$ is "approximately 1.08366...". Presumably, by this latter statement, Legendre meant that

$$\lim_{x \rightarrow \infty} A(x) = 1.08366.$$

It is precisely in regard to $A(x)$, where Legendre was in error, as we shall see below. In his memoir [9] of 1808, Legendre formulated another famous conjecture. Let k and l be integers which are relatively prime to one another. Then Legendre asserted that there exist infinitely many primes of the form $l + kn$ ($n = 0, 1, 2, 3, \dots$). In other words, if $\pi_{k,l}(x)$ denotes the number of primes p of the form $l + kn$ for which $p \leq x$, then Legendre conjectured that

$$(8) \quad \pi_{k,l}(x) \rightarrow \infty \quad \text{as} \quad x \rightarrow \infty.$$

Actually, the proof of (8) by Dirichlet in 1837 [2] provided several crucial ideas on how to approach the prime number theorem.

Although Legendre was the first person to publish a conjectural form of the prime number theorem, Gauss had already done extensive work on the theory of primes in 1792–3. Evidently Gauss considered the tabulation of primes as some sort of pastime and amused himself by compiling extensive tables on how the primes distribute themselves in various intervals of length 1000. We have included some of Gauss' tabulations as an Appendix. The first table, excerpted from [3, p. 436], covers the primes from 1 to 50,000. Each entry in the table represents an interval of length 1000. Thus, for example, there are 168 primes from 1 to 1000; 135 from 1001 to 2000; 127 from 3001 to 4000; and so forth. Gauss suspected that the density with which primes occurred in the neighborhood of the integer n was $1/\log n$, so that the number of primes in the interval $[a, b)$ should be approximately equal to

$$\int_a^b \frac{dx}{\log x}.$$

In the second set of tables, samples from [4, pp. 442–3], Gauss investigates the distribution of primes up to 3,000,000 and compares the number of primes found with the above integral. The agreement is striking. For example, between 2,600,000 and 2,700,000, Gauss found 6762 primes, whereas

$$\int_{2,600,000}^{2,700,000} \frac{dx}{\log x} = 6761.332.$$

Gauss never published his investigations on the distribution of primes. Nevertheless, there is little reason to doubt Gauss' claim that he first undertook his work in 1792–93, well before the memoir of Legendre was written. Indeed, there are several other known examples of results of the first rank which Gauss proved, but never communicated to anyone until years after the original work had been done. This was the case, for example, with the elliptic functions, where Gauss preceded Jacobi, and with Riemannian geometry, where Gauss anticipated Riemann. The only information beyond Gauss' tables concerning Gauss' work in the distribution of primes is contained in an 1849 letter to the astronomer Encke. We have included a translation of Gauss' letter.

In his letter Gauss describes his numerical experiments and his conjecture concerning $\pi(x)$. There are a number of remarkable features of Gauss' letter. On the second page of the letter, Gauss compares his approximation to $\pi(x)$, namely $\text{Li}(x)$, with Legendre's formula. The results are tabulated at the top of the second page and Gauss' formula yields a much larger numerical error. In a very prescient statement, Gauss defends his formula by noting that although Legendre's formula yields a smaller error, the rate of increase of Legendre's error term is much greater than his own. We shall see below that Gauss anticipated what is known as the "Riemann hypothesis." Another feature of Gauss' letter is that he casts doubt on Legendre's assertion about $A(x)$. He asserts that the numerical evidence does not support any conjecture about the limiting value of $A(x)$.

Gauss' calculations are awesome to contemplate, since they were done long before the days of high-speed computers. Gauss' persistence is most impressive. However, Gauss' tables are not error-free. My student, Edward Korn, has checked Gauss' tables using an electronic computer and has found a number of errors. We include the corrected entries in an appendix. In spite of these (remarkably few) errors, Gauss' calculations still provide overwhelming evidence in favor of the prime number theorem. Modern students of mathematics should take note of the great care with which data was compiled by such giants as Gauss. Conjectures in those days were rarely idle guesses. They were usually supported by piles of laboriously gathered evidence.

The next step toward a proof of the prime number theorem was a step in a completely different direction, and was taken by Dirichlet in 1837 [2]. In a beautiful memoir, Dirichlet proved Legendre's conjecture (8) concerning the infinitude of primes in an arithmetic progression. Dirichlet's work contained two radically new ideas, which we should discuss in some detail.

Let \mathbb{Z}_n denote the ring of residue classes modulo n , and let \mathbb{Z}_n^\times denote the group of units of \mathbb{Z}_n . Then \mathbb{Z}_n^\times is the so-called "group of reduced residue classes modulo n " and consists of those residue classes containing an element relatively prime to n . If k is an integer, let us denote by \bar{k} its residue class modulo n . Dirichlet's first brilliant idea was to introduce the **characters** of the group \mathbb{Z}_n^\times ; that is, the homomorphisms of \mathbb{Z}_n^\times into the multiplicative group \mathbb{C}^\times of non-zero complex numbers. If χ is such a character, then we may associate with χ a function (also denoted χ) from the semi-group \mathbb{Z}^* of non-zero integers as follows: Set

$$\begin{aligned}\chi(a) &= \chi(\bar{a}) \text{ if } (a, n) = 1 \\ &0 \text{ otherwise.}\end{aligned}$$

Then it is clear that $\chi: \mathbb{Z}^* \rightarrow \mathbb{C}^\times$ and has the following properties:

- (i) $\chi(a + n) = \chi(a)$,
- (ii) $\chi(aa') = \chi(a)\chi(a')$,
- (iii) $\chi(a) = 0$ if $(a, n) \neq 1$,
- (iv) $\chi(1) = 1$.

A function $\chi: \mathbb{Z}^* \rightarrow \mathbb{C}^\times$ satisfying (i)–(iv) is called a **numerical character** modulo n . Dirichlet's main result about such numerical characters was the so-called **orthogonality relations**, which assert the following:

$$\begin{aligned}\text{(A)} \quad \sum_a \chi(a) &= \phi(n) \text{ if } \chi \text{ is identically } 1, \\ &0 \text{ otherwise,}\end{aligned}$$

where a runs over a complete system of residues modulo n ;

$$(B) \quad \sum_x \chi(a) = \phi(n) \text{ if } a \equiv 1 \pmod{n}, \\ 0 \quad \text{otherwise,}$$

where χ runs over all numerical characters modulo n . Dirichlet's ideas gave birth to the modern theory of duality on locally compact abelian groups.

Dirichlet's second great idea was to associate to each numerical character modulo n and each real number $s > 1$, the following infinite series

$$(9) \quad L(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}.$$

It is clear that the series converges absolutely and represents a continuous function for $s > 1$. However, a more delicate analysis shows that the series (9) converges (although not absolutely) for $s > 0$ and represents a continuous function of s in this semi-infinite interval *provided that χ is not identically 1*. The function $L(s, \chi)$ has come to be called a **Dirichlet L-function**.

Note the following facts about $L(s, \chi)$: First $L(s, \chi)$ has a product formula of the form

$$(10) \quad L(s, \chi) = \prod_p \left(1 - \frac{\chi(p)}{p^s}\right)^{-1} \quad (s > 1),$$

where the product is taken over all primes p . The proof of (10) is very similar to the argument given above in Euler's proof of the infinity of prime numbers. Therefore, by (10),

$$(11) \quad \begin{aligned} \log L(s, \chi) &= - \sum_p \log \left(1 - \frac{\chi(p)}{p^s}\right) \\ &= - \sum_p \sum_{m=1}^{\infty} \frac{\chi(p^m)}{m p^{ms}}. \end{aligned}$$

Dirichlet's idea in proving the infinitude of primes in the arithmetic progression $a, a + n, a + 2n, \dots, (a, n) = 1$, was to imitate, somehow, Euler's proof of the infinitude of primes, by studying the function $L(s, \chi)$ for s near 1. The basic quantity to consider is

$$(12) \quad \begin{aligned} \sum_x \chi(a)^{-1} \log L(s, \chi) &= - \sum_p \sum_{m=1}^{\infty} \sum_x \frac{\chi(a)^{-1} \chi(p^m)}{m p^{ms}} \\ &= - \sum_p \sum_{m=1}^{\infty} \frac{1}{m p^{ms}} \chi(a)^{-1} \chi(p^m), \end{aligned}$$

where we have used (11). Let a^* be an integer such that $aa^* \equiv 1 \pmod{n}$. Then $\chi(a^*) = \chi(a)^{-1}$ by (i)-(iv). Moreover,

$$\begin{aligned}
 \sum_{\chi} \chi(a)^{-1} \chi(p^m) &= \sum_{\chi} \chi(a^* p^m) \\
 (13) \qquad \qquad \qquad &= \phi(n) \text{ if } a^* p^m \equiv 1 \pmod{n} \\
 &0 \text{ otherwise.}
 \end{aligned}$$

However, $a^* p^m \equiv 1 \pmod{n}$ is equivalent to $p^m \equiv a \pmod{n}$. Therefore, by (12) and (13), we have

$$(14) \qquad \sum_{\chi} \chi(a)^{-1} \log L(s, \chi) = -\phi(n) \sum_{\substack{p \\ p^m \equiv a \pmod{n}}} \sum_{m=1}^{\infty} \frac{1}{m p^{ms}}.$$

Thus, finally, we have

$$\begin{aligned}
 (15) \qquad -\frac{1}{\phi(n)} \sum_{\chi} \chi(a) \log L(s, \chi) &- \sum_{\substack{p \\ p^m \equiv a \pmod{n}}} \sum_{m=2}^{\infty} \frac{1}{m p^{mp}} \\
 &= \sum_{\substack{p \\ p \equiv a \pmod{n}}} \frac{1}{p^s} \quad (s > 1).
 \end{aligned}$$

From (15), we immediately see that in order to prove that there are infinitely many primes $p \equiv a \pmod{n}$, it is enough to show that the function

$$\sum_{p \equiv a \pmod{n}} \frac{1}{p^s}$$

tends to $+\infty$ as s approaches 1 from the right. But it is fairly easy to see that as $s \rightarrow 1+$, the sum

$$\sum_{\substack{p \\ p^m \equiv a \pmod{n}}} \sum_{m=2}^{\infty} \frac{1}{m p^{ms}}$$

remains bounded. Thus, it suffices to show that

$$-\frac{1}{\phi(n)} \sum_{\chi} \chi(a)^{-1} \log L(s, \chi) \rightarrow +\infty \quad (s \rightarrow 1+).$$

However, if χ_0 denotes the character which is identically 1, then it is easy to see that

$$-\frac{1}{\phi(n)} \chi_0(a)^{-1} L(s, \chi_0) \rightarrow +\infty \text{ as } s \rightarrow 1+.$$

Therefore, it is enough to show that if $\chi \neq \chi_0$, then $\log L(s, \chi)$ remains bounded as $s \rightarrow 1+$. We have already mentioned that $L(s, \chi)$ is continuous for $s > 0$ if $\chi \neq \chi_0$. Therefore, it suffices to show that $L(1, \chi) \neq 0$. And this is precisely what Dirichlet showed.

Dirichlet's theorem on primes in arithmetic progressions was one of the major achievements of 19th century mathematics, because it introduced a fertile new idea into number theory—that analytic methods (in this case the study of the Dirichlet L -series) could be fruitfully applied to arithmetic problems (in this case the problem of primes in arithmetic progressions). To the novice, such an application of analysis to number theory would seem to be a waste of time. After all, number theory is the study of the discrete, whereas analysis is the study of the continuous; and what should one have to do with the other! However, Dirichlet's 1837 paper was the beginning of a revolution in number-theoretic thought, the substance of which was to apply analysis to number theory. At first, undoubtedly, mathematicians were very uncomfortable with Dirichlet's ideas. They regarded them as very clever devices, which would eventually be supplanted by completely arithmetic ideas. For although analysis might be useful in proving results about the integers, surely the analytic tools were not intrinsic. Rather, they entered the theory of the integers in an inessential way and could be eliminated by the use of suitably sophisticated arithmetic. However, the history of number theory in the 19th century shows that this idea was eventually repudiated and the rightful connection between analysis and number theory came to be recognized.

The first major progress toward a proof of the prime number theorem after Dirichlet was due to the Russian mathematician Tchebycheff in two memoirs [12, 13] written in 1851 and 1852. Tchebycheff introduced the following two functions of a real variable x :

$$\theta(x) = \sum_{p \leq x} \log p,$$

$$\psi(x) = \sum_{p^m \leq x} \log p,$$

where p runs over primes and m over positive integers. Tchebycheff proved that the prime number theorem (1) is equivalent to either of the two statements

$$(16) \quad \lim_{x \rightarrow \infty} \frac{\theta(x)}{x} = 1,$$

$$(17) \quad \lim_{x \rightarrow \infty} \frac{\psi(x)}{x} = 1.$$

Moreover, Tchebycheff proved that if $\lim_{x \rightarrow \infty} (\theta(x)/x)$ exists, then its value must be 1. Furthermore, Tchebycheff proved that

$$(18) \quad .92129 \leq \liminf_{x \rightarrow \infty} \frac{\pi(x)}{x/\log x} \leq 1 \leq \limsup_{x \rightarrow \infty} \frac{\pi(x)}{x/\log x} \leq 1.10555.$$

Tchebycheff's methods were of an elementary, combinatorial nature, and as such were not powerful enough to prove the prime number theorem.

The first giant strides toward a proof of the prime number theory were taken by B. Riemann in a memoir [10] written in 1860. Riemann followed Dirichlet in connecting problems of an arithmetic nature with the properties of a function of a continuous variable. However, where Dirichlet considered the functions $L(s, \chi)$ as functions of a real variable s , Riemann took the decisive step in connecting arithmetic with the theory of functions of a complex variable. Riemann introduced the following function:

$$(19) \quad \zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s},$$

which has come to be known as the **Riemann zeta function**. It is reasonably easy to see that the series (19) converges absolutely and uniformly for s in a compact subset of the half-plane $\operatorname{Re}(s) > 1$. Thus, $\zeta(s)$ is analytic for $\operatorname{Re}(s) > 1$. Moreover, by using the same sort of argument used in Euler's proof of the infinitude of primes, it is easy to prove that

$$(20) \quad \zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1} \quad (\operatorname{Re}(s) > 1),$$

where the product is extended over all primes p . Euler's proof of the infinitude of primes suggests that the behavior of $\zeta(s)$ for $s = 1$ is somehow connected with the distribution of primes. And, indeed, this is the case.

Riemann proved that $\zeta(s)$ can be analytically continued to a function which is meromorphic in the whole s -plane. The only singularity of $\zeta(s)$ occurs at $s = 1$ and the Laurent series about $s = 1$ looks like

$$(21) \quad \zeta(s) = \frac{1}{s-1} + a_0 + a_1(s-1) + \cdots.$$

Moreover, if we set

$$(22) \quad R(s) = s(s-1)\pi^{-s/2}\Gamma(s/2)\zeta(s),$$

then $R(s)$ is an entire function of s and satisfies the functional equation

$$(23) \quad R(s) = R(1-s).$$

To see the immediate connection between the Riemann zeta function and the distribution of primes, let us return to Euler's proof of the infinitude of primes. A variation on the idea of Euler's proof is as follows: Suppose that there were only finitely many primes p_1, \dots, p_N . Then by (20), $\zeta(s)$ would be bounded as s tends to 1, which contradicts equation (21). Thus, the presence of a pole of $\zeta(s)$ at $s = 1$ immediately implies that there are infinitely many primes. But the connection between the zeta function and the distribution of primes runs even deeper.

Let us consider the following heuristic argument: From equation (20), it is easy to deduce that

$$(24) \quad \frac{\zeta'(s)}{\zeta(s)} = \sum_p \sum_{m=1}^{\infty} (\log p) p^{-ms} \quad (\operatorname{Re}(s) > 1).$$

Moreover, by residue calculus, it is easy to verify that

$$(25) \quad \lim_{T \rightarrow \infty} \frac{1}{2m} \int_{2-iT}^{2+iT} \frac{a^s}{s} ds = \begin{cases} 1, & x < 1 \\ 0, & x > 1. \end{cases}$$

Therefore, assuming that interchange of limit and summation is justified, we see that for x not equal to an integer, we have

$$(26) \quad \begin{aligned} \lim_{T \rightarrow \infty} \frac{1}{2m} \int_{2-iT}^{2+iT} \frac{x^s}{s} \frac{\zeta'(s)}{\zeta(s)} ds &= \sum_p \sum_{m=1}^{\infty} (\log p) \lim_{T \rightarrow \infty} \frac{1}{2\pi i} \int_{2-iT}^{2+iT} \left(\frac{x}{p^m}\right)^s \frac{1}{s} ds \\ &= \sum_{p^m \leq x} \log p \quad (\text{by equation (25)}) \\ &= \psi(x). \end{aligned}$$

Thus, we see that there is an intimate connection between the function $\psi(x)$ and $\zeta(s)$. This connection was first exploited by Riemann, in his 1860 paper.

Note that the function

$$(27) \quad \frac{x^s}{s} \frac{\zeta'(s)}{\zeta(s)}$$

has poles at $s = 0$ and at all zeroes ρ of $\zeta(s)$. Moreover, note that by equation (20), we see that $\zeta(s) \neq 0$ for $\operatorname{Re}(s) > 1$. Therefore, all zeroes of $\zeta(s)$ lie in the half-plane $\operatorname{Re}(s) \leq 1$. Further, since $R(s)$ is entire and $\zeta(s) \neq 0$ for $\operatorname{Re}(s) > 1$, the functional equation (23) implies that the only zeroes of $\zeta(s)$ for which $\operatorname{Re}(s) < 0$ are at $s = -2, -4, -6, -8, \dots$, and these are all simple zeroes and are called the **trivial zeroes** of $\zeta(s)$. Thus, we have shown that all non-trivial zeroes of $\zeta(s)$ lie in the strip $0 \leq \operatorname{Re}(s) \leq 1$. This strip is called the **critical strip**. The residue of (27) at a non-trivial zero ρ is

$$\frac{x^\rho}{\rho}.$$

Thus, if σ is a large negative number, and if $C_{\sigma, T}$ denotes the rectangle with vertices $\sigma \pm iT, 2 \pm iT$, then Cauchy's theorem implies that

$$(28) \quad \frac{1}{2\pi i} \int_{2-iT}^{2+iT} \frac{x^s}{s} \frac{\zeta'(s)}{\zeta(s)} ds = \frac{1}{2\pi i} \left[\int_{\sigma-iT}^{\sigma+iT} + \int_{\sigma+iT}^{2+iT} + \int_{2+iT}^{2-iT} \right] \frac{x^s}{s} \frac{\zeta'(s)}{\zeta(s)} ds + R(\sigma, T),$$

where $R(\sigma, T)$ denotes the sum of the residues of the function (27) at the poles inside

$C_{\sigma,T}$. By letting σ and T tend to infinity and by applying equations (26) and (28), Riemann arrived at the following remarkable formula, known as **Riemann's explicit formula**

$$(29) \quad \psi(x) = x - \sum_{\rho} \frac{x^{\rho}}{\rho} - \frac{\zeta'(0)}{\zeta(0)} - \frac{1}{2} \log(1 - x^{-2}),$$

where ρ runs over all non-trivial zeroes of the Riemann zeta function. Riemann's formula is surprising for at least two reasons. First, it connects the function $\psi(x)$, which is connected with the distribution of primes, with the distribution of the zeroes of the Riemann zeta function. That there should be any connection at all is amazing. But, secondly, the formula (29) explicitly puts in evidence a form of the prime number theorem by equating $\psi(x)$ with x plus an error term which depends on the zeroes of the zeta function. If we denote this error term by $E(x)$, then we see that the prime number theorem is equivalent to the assertion

$$(30) \quad \lim_{x \rightarrow \infty} \frac{E(x)}{x} = 0,$$

which, in turn, is equivalent to the assertion

$$(31) \quad \lim_{x \rightarrow \infty} \frac{1}{x} \sum_{\rho} \frac{x^{\rho}}{\rho} = 0.$$

Riemann was unable to prove (31), but he made a number of conjectures concerning the distributions of the zeroes ρ from which the statement (31) follows immediately. The most famous of Riemann's conjectures is the so-called **Riemann hypothesis**, which asserts that all non-trivial zeroes of $\zeta(s)$ lie on the line $\operatorname{Re}(s) = \frac{1}{2}$, which is the line of symmetry of the functional equation (23). This conjecture has resisted all attempts to prove it for more than a century and is one of the most celebrated open problems in all of mathematics. However, if the Riemann hypothesis is true, then

$$\left| \frac{x^{\rho}}{\rho} \right| = x^{\frac{1}{2}} \frac{1}{|\rho|}$$

and from this fact and equation (29), it is possible to prove that

$$(32) \quad \psi(x) = x + O(x^{\frac{1}{2} + \varepsilon})$$

for every $\varepsilon > 0$, where $O(x^{\frac{1}{2} + \varepsilon})$ denotes a function $f(x)$ such that $f(x)/x^{\frac{1}{2} + \varepsilon}$ is bounded for all large x . Thus, the Riemann hypothesis implies (31) in a trivial way, and hence the prime number theorem follows from the Riemann hypothesis. What is perhaps more striking is the fact that *if (32) holds then the Riemann hypothesis is true*. Thus, the prime number theorem in the sharp form (32) is equivalent to the Riemann hypothesis. We see, therefore, that the connection between the zeta function and the

distribution of primes is no accidental affair, but somehow is woven into the fabric of nature.

In his memoir, Riemann made many other conjectures. For example, if $N(T)$ denotes the number of non-trivial zeroes ρ of $\zeta(s)$ such that $-T \leq \text{Im}(\rho) \leq T$, then Riemann conjectured that

$$(33) \quad N(T) = \frac{1}{2\pi} T \log T - \frac{1 + \log(2\pi)}{2\pi} T + O(\log T).$$

The formula (33) was first proven by von-Mangoldt in 1895 [14]. An interesting line of research has been involved in obtaining estimates for the number of non-trivial zeroes ρ on the line $\text{Re}(s) = \frac{1}{2}$. Let $M(T)$ denote the number of ρ such that $\text{Re}(s) = \frac{1}{2}$, $-T \leq \text{Im}(s) \leq T$. Then Hardy [6] in 1912, proved that $M(T)$ tends to infinity as T tends to infinity. Later, Hardy [7] improved his argument to prove that $M(T) > AT$, where A is a positive constant, not depending on T . The ultimate result of this sort was obtained by Atle Selberg in 1943 [11]. He proved that $M(T) > AT \log T$ for some positive constant A . In view of equation (33), Selberg's result shows that a positive percentage of the zeroes of $\zeta(s)$ actually lie on the line $\text{Re}(s) = \frac{1}{2}$. This result represents the best progress made to date in attempting to prove the Riemann hypothesis.

Fortunately, it is not necessary to prove the Riemann hypothesis in order to prove the prime number theorem in the form (17). However, it is necessary to obtain some information about the distribution of the zeroes of $\zeta(s)$. Such information was obtained independently by Hadamard [5] and de la Vallée Poussin [1] in 1896, thereby providing the first complete proofs of the prime number theorem. Although their proofs differ in detail, they both establish the existence of a zero-free region for $\zeta(s)$, the existence of which serves as a substitute for the Riemann hypotheses in the reasoning presented above. More specifically, they proved that there exist constants a, t_0 such that $\zeta(\sigma + it) \neq 0$ if $\sigma \geq 1 - 1/a \log |t|$, $|t| \geq t_0$. This zero-free region allows one to prove the prime number theorem in the form

$$(34) \quad \psi(x) = x + O(xe^{-c(\log x)^{1/4}}).$$

Please note, however, that the error term in (34) is much larger than the error term predicted by the Riemann hypothesis.

Thus, the prime number theorem was finally proved after a century of hard work by many of the world's best mathematicians. It is grossly unfair to attribute proof of such a theorem to the genius of a single individual. For, as we have seen, each step in the direction of a proof was conditioned historically by the work of preceding generations. On the other hand, to deny that there is genius in the work which led up to the ultimate proof would be equally unfair. For at each step in the chain of discovery, brilliant and fertile ideas were discovered, and provided the material out of which to fashion the next link.

APPENDIX A: Samples from Gauss' Tables. TABLE 1 (*Werke, II*, p. 436)

1	168	26	98
2	135	27	101
3	127	28	94
4	120	29	98
5	119	30	92
6	114	31	95
7	117	32	92
8	107	33	106
9	110	34	100
10	112	35	94
11	106	36	92
12	103	37	99
13	109	38	94
14	105	39	90
15	102	40	96
16	108	41	88
17	98	42	101
18	104	43	102
19	94	44	85
20	102	45	96
21	98	46	86
22	104	47	90
23	100	48	95
24	104	49	89
25	94	50	98

The frequency of primes. TABLE 2 (*Werke, II*, p. 443) 2000000...3000000

	210	220	230	240	250	260	270	280	290	300	
0							1				1
1	3	2	2	4	1	3	4	2	2	2	25
2	10	9	9	11	9	5	10	7	15	13	98
3	32	27	29	32	37	35	28	43	30	44	337
4	69	69	73	86	78	88	71	95	85	64	778
5	119	146	138	136	147	136	158	135	140	153	1408
6	197	183	179	176	193	194	195	195	179	187	1878
7	204	201	205	194	189	180	201	188	222	214	1998
8	157	168	168	158	151	170	142	145	132	134	1525
9	115	109	113	112	102	88	96	87	109	103	1034
10	63	52	44	55	58	58	53	67	53	58	561
11	21	18	30	28	23	24	22	24	18	15	223
12	8	9	10	7	7	13	17	9	8	11	99
13	2	4		1	5	6	1	2	5	1	27
14		3					1		2		6
15										1	1
16											
17								1			1
	6874	6857	6849	6787	6766	6804	6762	6714	6744	6705	6862

APPENDIX B: Gauss' Letter to Enke.

My distinguished friend:

Your remarks concerning the frequency of primes were of interest to me in more ways than one. You have reminded me of my own endeavors in this field which began in the very distant past, in 1792 or 1793, after I had acquired the Lambert supplements to the logarithmic tables. Even before I had begun my more detailed investigations into higher arithmetic, one of my first projects was to turn my attention to the decreasing frequency of primes, to which end I counted the primes in several chiliads (*intervals of length 1000*; *Trans.*) and recorded the results on the attached white pages. I soon recognized that behind all of its fluctuations, this frequency is on the average inversely proportional to the logarithm, so that the number of primes below a given bound n is approximately equal to

$$\int \frac{dn}{\log n},$$

where the logarithm is understood to be hyperbolic. Later on, when I became acquainted with the list in Vega's tables (1796) going up to 400031, I extended my computation further, confirming that estimate. In 1811, the appearance of Chernau's cribrum gave me much pleasure and I have frequently (since I lack the patience for a continuous count) spent an idle quarter of an hour to count another chiliad here and there; although I eventually gave it up without quite getting through a million. Only some time later did I make use of the diligence of Goldschmidt to fill some of the remaining gaps in the first million and to continue the computation according to Burkhardt's tables. Thus (for many years now) the first three million have been counted and checked against the integral. A small excerpt follows:

TABLE A

Below	Here are Prime	Integral $\int \frac{dn}{\log n}$	Error	Your Formula	Error
500000	41556	41606.4	+ 50.4	41596.9	+ 40.9
1000000	78501	79627.5	+ 126.5	78672.7	+ 171.7
1500000	114112	114263.1	+ 151.1	114374.0	+ 264.0
2000000	148883	149054.8	+ 171.8	149233.0	+ 350.0
2500000	183016	183245.0	+ 229.0	183495.1	+ 479.1
3000000	216745	216970.6	+ 225.6	217308.5	+ 563.5

I was not aware that Legendre had also worked on this subject; your letter caused me to look in his *Théorie des Nombres*, and in the second edition I found a few pages on the subject which I must have previously overlooked (or, by now, forgotten). Legendre used the formula

$$\frac{n}{\log n - A},$$

where A is a constant which he sets equal to 1.08366. After a hasty computation, I find in the above cases the deviations

TABLE B

— 23,3
+ 42,2
+ 68,1
+ 92,8
+159,1
+167,6

These differences are even smaller than those from the integral, but they seem to grow faster with n so that it is quite possible they may surpass them. To make the count and the formula agree, one would have to use, respectively, instead of $A = 1.08366$, the following numbers:

TABLE C

1,09040
1,07682
1,07582
1,07529
1,07179
1,07297

It appears that, with increasing n , the (average) value of A decreases; however, I dare not conjecture whether the limit as n approaches infinity is 1 or a number different from 1. I cannot say that there is any justification for expecting a very simple limiting value; on the other hand, the excess of A over 1 might well be a quantity of the order of $1/\log n$. I would be inclined to believe that the differential of the function must be simpler than the function itself.

If $dn/\log n$ is postulated for the function, Legendre's formula would suggest that the differential function might be something of the form $dn/(\log n - (A-1))$. By the way, for large n , your formula could be considered to coincide with

$$\frac{n}{\log n - (1/2k)},$$

where k is the modulus of Brigg's logarithms; that is, with Legendre's formula, if we put $A = 1/2k = 1.1513$.

Finally, I want to remark that I noticed a couple of disagreements between your counts and mine.

Between	59000	and	60000,	you have	95,	while I have	94
	101000		102000		94		93.

The first difference possibly results from the fact that, in Lambert's Supplement, the prime 59023 occurs twice. The chiliad from 101000 — 102000 in Lambert's Supplement is virtually crawling with errors; in my copy, I have indicated seven numbers which are not primes at all, and supplied two missing ones. Would it not be possible to induce young Mr. Dase to count the primes in the following (few) millions, using the tables at the Academy which, I am afraid, are not intended for public distribution? In this case, let me remark that in the 2nd and 3rd million, the count is, according to my instructions, based on a special scheme which I myself have employed in counting the first million. The counts for each 100000 are indicated on a single page in 10 columns, each column belonging to one myriad (*an interval of length 10000; Trans.*); an additional column in front (left) and another column following it on the right; for example here is a vertical column and the two additional columns for the interval 10000000 to 11000000 — — —

As an illustration, take the first vertical column. In the myriad 1000000 to 1010000 there are 100 Hecatontades; (*intervals of length 100; Trans.*) among them one containing a single prime, none containing two or three primes; two containing four each; eleven containing 5 each, etc., yielding altogether $752 = 1.1 + 4.2 + 5.11 + 6.14 + \dots$ primes. The last column contains the totals from the other ten. The numbers 14, 15, 16 in the first vertical column are superfluous since no hecatontades occur containing that many primes; but on the following pages they are needed. Finally the 10 pages are again combined into one and thus comprise the entire second million.

It is high time to quit — — — . With most cordial wishes for your good health

Yours, as ever,
C. F. Gauss

Göttingen, 24 December 1849.

APPENDIX C: Corrections to Gauss' Tables

THOUSANDS	GAUSS	ACTUAL	Δ
20	102	104	-2
159	87	77	+10
199	96	86	+10
206	85	83	+2
245	78	88	-10
289	85	77	+8
290	84	85	-1
334	80	81	-1
352	80	81	-1
354	79	76	+3
500	UP	TO HERE	+18
TOTALS			Δ
500,000	41,556	41,538	+18
1,000,000	78,501	78,498*	+3
1,500,000	114,112	114,156*	-44
2,000,000	148,883	148,934*	-51
2,500,000	183,016	183,073*	-57
3,000,000	216,745	216,817*	-72

* from *List of Prime Numbers from 1 to 10,006,771*, by D. N. Lehmer, (adjusted: He counts 1 as a prime).

Research supported by NSF Grant GP 31280X. This article was presented to the History of Mathematics Seminar at the University of Maryland on March 20, 1972. The author wishes to thank Professor Gertrude Ehrlich for preparing the translation of Gauss' letter which appears in Appendix B. He also wishes to thank Mr. Edward Korn for providing the calculations of Appendix C.

References

1. Ch. de la Vallée Poussin, *Recherches analytiques sur la théorie des nombres premiers. Première partie. La fonction $\zeta(s)$ de Riemann et les nombres premiers en général. Deuxième partie: Les fonc-*

tions de Dirichlet et les nombres premiers de la forme linéaire $Mx+N$. Troisième partie: Les formes quadratiques de déterminant négatif, Ann. Soc. Sci. Bruxelles, 20 (1896) 183–256, 281–397.

2. L. Dirichlet, Über den Satz: das jede unbegrenzte arithmetische Progression, deren erstes Glied und Differenz keinen gemeinschaftlichen Factor sind, unendlichen viele Primzahlen enthält, 1837; Mathematische Abhandlungen, Bd. 1, (1889) 313–342.

3. C. F. Gauss, Tafel der Frequenz der Primzahlen, Werke, II (1872) 436–442.

4. ———, Gauss an Encke, Werke, II (1872) 444–447.

5. J. Hadamard, Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques, Bull. Soc. Math. de France, 24 (1896) 199–220.

6. G. H. Hardy, Sur les zéros de la fonction $\zeta(s)$ de Riemann, Comptes Rendus, 158 (1914) 1012–14.

7. ———, and J. E. Littlewood, The zeros of Riemann's zeta function on the critical line, Math. Zeit., 10 (1921) 283–317.

8. A. M. Legendre, Essai sur la théorie de Nombres, 1st ed., 1798, Paris, p. 19.

9. ———, Essai sur la Théorie de Nombres, 2nd ed. 1808, Paris, p. 394.

10. B. Riemann, Über die Anzahl der Primzahlen unter einer gegebenen Grösse, Gesammelte Mathematische Werke, 2nd Aufl., (1892) 145–155.

11. A. Selberg, On the zeros of Riemann's zeta function, Skr. Norske Vid. Akad., Oslo (1942) no. 10.

12. P. Tchebycheff, Sur la fonction qui détermine la totalité de nombres premiers inférieurs à une limite donnée, Oeuvres, 1 (1899) 27–48.

13. ———, Mémoire sur les nombres premiers, Oeuvres, 1 (1899) 49–70.

14. H. von Mangoldt, Auszug aus einer Arbeit unter dem Titel: Zu Riemann's Abhandlung über die Anzahl der Primzahlen unter einer gegebenen Grösse, Sitz. König. Preus. Akad. Wiss. zu Berlin, (1894) 337–350, 883–895.

DIFFERENTIATION UNDER THE INTEGRAL SIGN*

HARLEY FLANDERS, Tel-Aviv University

1. Introduction. Everyone knows the Leibniz rule for differentiating an integral:

$$(1.1) \quad \begin{aligned} & \frac{d}{dt} \left(\int_{g(t)}^{h(t)} F(x, t) dx \right) \\ &= \left\{ F[h(t), t]h'(t) - F[g(t), t]g'(t) \right\} + \int_{g(t)}^{h(t)} \frac{\partial F(x, t)}{\partial t} dx. \end{aligned}$$

We are all fond of this formula, although it is seldom if ever used in such generality. Usually, either the limits are constants, or the integrand is independent of the time t . Frequent cases are

$$\frac{d}{dt} \int_a^t F(x) dx = F(t), \quad \frac{d}{dt} \int_0^\infty F(x, t) dx = \int_0^\infty \frac{\partial F(x, t)}{\partial t} dx.$$

* Presented May 5, 1972 to the Rocky Mountain Section meeting, Southern Colorado State College, Pueblo, CO.

One proof runs as follows, modulo precisely stated hypotheses and some analytic details. Set

$$(1.2) \quad \Phi(u, v, t) = \int_u^v F(x, t) dx,$$

$u = g(t)$, and $v = h(t)$. By the chain rule

$$\frac{d}{dt} \Phi[g(t), h(t), t] = \left(\frac{\partial \Phi}{\partial u} \dot{g} + \frac{\partial \Phi}{\partial v} \dot{h} \right) + \frac{\partial \Phi}{\partial t}.$$

The first two terms are bracketed because they measure all changes due to variation of the interval of integration $[g(t), h(t)]$, and they are evaluated by applying the Fundamental Theorem to (1.2). The third term measures change due to variation of the integrand. If enough smoothness is assumed to justify interchange of the integration and differentiation operators, then

$$(1.3) \quad \frac{\partial \Phi}{\partial t} = \frac{\partial}{\partial t} \int_u^v F(x, t) dx = \int_u^v \frac{\partial F(x, t)}{\partial t} dx.$$

We shall discuss generalizations of the Leibniz rule to more than one dimension. Such generalizations seem to be common knowledge among physicists, some differential geometers, and applied mathematicians who work in continuum mechanics, but are virtually unheard of among most mathematicians. I cannot find a single mention of such formulas in the current advanced calculus and several variable texts, except for Loomis and Sternberg [4].

REMARK: A nice approach to (1.3) is via interchange of the order of integration (Fubini's Theorem):

$$\begin{aligned} \int_u^v \frac{\partial F(x, t)}{\partial t} dx &= \frac{d}{dt} \int_a^t ds \int_u^v \frac{\partial F(x, s)}{\partial s} dx \\ &= \frac{d}{dt} \int_u^v dx \int_a^t \frac{\partial F(x, s)}{\partial s} ds \\ &= \frac{d}{dt} \int_u^v [F(x, t) - F(x, a)] dx \\ &= \frac{d}{dt} \int_u^v F(x, t) dx. \end{aligned}$$

See for example Fleming [3] for details.

2. Another proof. We shall concentrate on the change due to variation of the interval. This puts us in the proper frame of mind for generalization to more dimensions, where the real difficulties are with the moving domain, not with the time-

varying integrand. Anyhow, we know how to separate the domain variation from the integrand variation by the chain rule device used above.

Thus we are concentrating on

$$\frac{d}{dt} \int_{g(t)}^{h(t)} F(x) dx.$$

The domain of integration, the interval $C_t = [g(t), h(t)]$ is moving with time, but we have no idea how points interior to the domain move. Only the motion of the boundary points has been prescribed; no one said anything about interior points!

Even though we know in advance that the answer is independent of how the interior points may move, we shall stubbornly insist that they have a definite motion.

Imagine the interval C_t is a worm crawling along the x -axis. As it stretches and shrinks and does the things worms do, each point of its body follows some irregular trajectory. Suppose initially the worm's points are labeled u , where $a \leq u \leq b$, and at time t , the point initially at u is at $x = x(u, t)$. Now, a worm can only shrink so much, so $\partial x / \partial u > 0$. For each t , the map $u \rightarrow x(u, t)$ is smooth one-one with smooth (continuously differentiable) inverse. We might write ϕ_t for this map at t :

$$\phi_t(u) = x(u, t),$$

$$\phi_t: [a, b] \rightarrow [\phi_t(a), \phi_t(b)] = [g(t), h(t)] = C_t.$$

By the formula for change of variable in a simple integral,

$$\int_{g(t)}^{h(t)} F(x) dx = \int_{\phi_t(a)}^{\phi_t(b)} F(x) dx = \int_a^b F[x(u, t)] \frac{\partial x}{\partial u} du.$$

This transition is excellent, because it has changed the integral over a moving domain to one over a fixed domain. We pay for this fixed domain with a time-varying integrand. No matter, we like it; we thrive on differentiation under the integral sign:

$$\begin{aligned} \frac{d}{dt} \int_{g(t)}^{h(t)} F(x) dx &= \frac{d}{dt} \int_a^b F[x(u, t)] \frac{\partial x}{\partial u} du \\ &= \int_a^b \frac{\partial}{\partial t} \left\{ F[x(u, t)] \frac{\partial x}{\partial u} \right\} du \\ &= \int_a^b \left\{ F'[x(u, t)] \frac{\partial x}{\partial t} \frac{\partial x}{\partial u} + F[x(u, t)] \frac{\partial^2 x}{\partial u \partial t} \right\} du. \end{aligned}$$

The fixed domain has done its job, and we return to the moving domain. The instantaneous velocity is $v = v(u, t) = \partial x / \partial t$, which we also consider as a function of x and t via the transformation $(u, t) \leftrightarrow (x, t)$. When t is fixed,

$$\frac{\partial^2 x}{\partial u \partial t} = \frac{\partial v}{\partial u} = \left(\frac{\partial v}{\partial u} \middle/ \frac{\partial x}{\partial u} \right) \frac{\partial x}{\partial u} = \frac{\partial v}{\partial x} \frac{\partial x}{\partial u},$$

hence

$$\begin{aligned} \frac{d}{dt} \int &= \int_{\phi_t(a)}^{\phi_t(b)} \left[F'(x)v + F(x) \frac{\partial v}{\partial x} \right] dx \\ &= \int_{\phi_t(a)}^{\phi_t(b)} \frac{\partial}{\partial x} [F(x)v] dx = \int_{g(x)}^{h(x)} \frac{\partial}{\partial x} [F(x)v] dx, \end{aligned}$$

by the change of variable formula in reverse gear. Note that the time t is fixed in this process; the whole integration takes place instantaneously.

We pause momentarily to inspect our progress. The derivative in question has been expressed as an integral over the moving domain. The integrand depends on the velocity v at each point of the domain, but it just happens that the integrand is an exact derivative, so the answer depends only on the boundary values. At the boundary points $g(t)$ and $f(t)$, the velocities are $\dot{g}(t)$ and $\dot{f}(t)$ respectively, so finally

$$\frac{d}{dt} \int_{g(t)}^{h(t)} F(x) dx = F[h(t)]\dot{h}(t) - F[g(t)]\dot{g}(t).$$

This might seem a silly approach to the problem because (1) it introduces an unnecessary quantity v , and (2) it evades using the fundamental theorem initially, only to use it in the end after all. Yet there is an essential idea here, reduction to a fixed domain, and it wins the day when we generalize.

3. A plane formula. Imagine a moving domain D_t in the x, y -plane (Fig. 1).

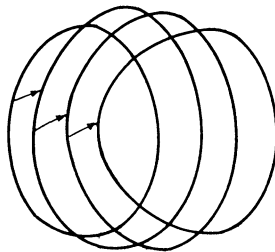


FIG. 1

We are also given a function $F(x, y, t)$. The problem is to find

$$\frac{d}{dt} \iint_{D_t} F(x, y, t) dx dy.$$

Already the ugly method of the last section is looking better, because it is not immediately clear what replaces the two terms in (1.1) that resulted one way or the other from use of the fundamental theorem. Actually, on second thought, the fundamental theorem just may prove relevant, but in its two dimensional form, viz., Green's Theorem.

Certainly our first move should be separation of boundary variation from integrand variation. This is easy enough by the chain rule device in the first section and results in

$$(3.1) \quad \begin{aligned} \frac{d}{dt} \iint_{D_t} F(x, y, t) dx dy \Big|_{t=t_0} \\ = \frac{d}{dt} \iint_{D_t} F(x, y, t_0) dx dy \Big|_{t=t_0} + \iint_{D_{t_0}} \frac{\partial F}{\partial t} \Big|_{t=t_0} dx dy. \end{aligned}$$

This is routine. The essence of the problem is to find

$$\frac{d}{dt} \iint_{D_t} F(x, y) dx dy.$$

This we shall do by a physicist's argument.

Look at two successive domains D_t and D_{t+dt} . See Fig. 2.

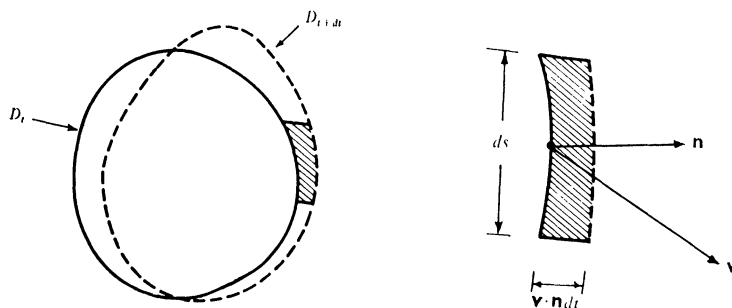


FIG. 2.

Let $\mathbf{v} = \mathbf{v}(x, y, t)$ denote the velocity vector at a boundary point (x, y) of D_t and let \mathbf{n} denote the outward unit normal. In the difference

$$\iint_{D_{t+dt}} F(x, y) dx dy - \iint_{D_t} F(x, y) dx dy,$$

everything in the overlap of D_t and D_{t+dt} cancels; only the thin boundary strip makes a contribution. From the detail, this contribution is

$$F(x, y) (\mathbf{v} dt) \cdot (\mathbf{n} ds)$$

up to higher order differentials, where ds is the element of arc length. (Disclaimer: I said it's a physicist's proof!) Hence

$$\frac{1}{dt} \left(\iint_{D_{t+dt}} - \iint_{D_t} \right) \approx \int_{\partial D_t} F(x, y) \mathbf{v} \cdot \mathbf{n} ds,$$

where ∂ denotes boundary. Before taking limits, we compute $\mathbf{v} \cdot \mathbf{n} ds$. We rotate the

unit tangent $(dx/ds, dy/ds)$ backwards through a right angle to obtain $\mathbf{n} = (dy/ds, -dx/ds)$, hence

$$\mathbf{v} \cdot \mathbf{n} ds = (u, v) \cdot (dy, -dx) = u dy - v dx.$$

Therefore

$$(3.2) \quad \frac{d}{dt} \iint_{D_t} F(x, y) dx dy = \int_{\partial D_t} F(x, y) (u dy - v dx).$$

We can transform the boundary integral into an integral over D_t by Green's Theorem. Let us do this and also combine (3.1) and (3.2) for the result of this section, a Leibniz rule in the plane:

$$(3.3) \quad \begin{aligned} \frac{d}{dt} \iint_{D_t} F(x, y, t) dx dy &= \int_{\partial D_t} F(u dy - v dx) + \iint_{D_t} \frac{\partial F}{\partial t} dx dy \\ &= \iint_{D_t} \left[\operatorname{div}(F\mathbf{v}) + \frac{\partial F}{\partial t} \right] dx dy. \end{aligned}$$

Here

$$\operatorname{div}(F\mathbf{v}) = \frac{\partial}{\partial x}(Fu) + \frac{\partial}{\partial y}(Fv) = (\operatorname{grad} F) \cdot \mathbf{v} + F \operatorname{div} \mathbf{v}.$$

4. A space formula. Consider a fluid flowing through a region of space. The Lagrange (historical) description of the flow gives the position $\mathbf{x} = \mathbf{x}(\mathbf{u}, t)$ at time t of the particle of fluid originally at point \mathbf{u} . The Euler description gives the velocity $\mathbf{v} = \mathbf{v}(\mathbf{x}, t)$ at present time t of the particle now at position \mathbf{x} . Suppose we are given a domain D_t that moves with the flow. Suppose also we are given a function $F(\mathbf{x}, t)$ on the region of flow. The following formula, with a physicist's proof, can be found in Prager [5], or Sokolnikoff and Redheffer [6].

$$(4.1) \quad \begin{aligned} \frac{d}{dt} \iiint_{D_t} F(\mathbf{x}, t) dx dy dz &= \iint_{\partial D_t} F\mathbf{v} \cdot d\boldsymbol{\sigma} + \iiint_{D_t} \frac{\partial F}{\partial t} dx dy dz \\ &= \iiint_{D_t} \left[\operatorname{div}(F\mathbf{v}) + \frac{\partial F}{\partial t} \right] dx dy dz. \end{aligned}$$

Here $d\boldsymbol{\sigma}$ is the vectorial area element on the closed surface ∂D_t , so that

$$d\boldsymbol{\sigma} = (dy dz, dz dx, dx dy) = \mathbf{n} d\sigma,$$

where \mathbf{n} is the outward unit normal and $d\sigma$ is the element of area. We shall give a mathematical proof of (4.1), without worrying much about minimal smoothness conditions. Note that the two versions of the formula are equivalent by Gauss' divergence theorem.

We shall use index notation for coordinates. The initial position is $\mathbf{u} = (u^1, u^2, u^3)$, the moving point is $\mathbf{x} = (x^1, x^2, x^3)$, and the velocity is $\mathbf{v} = (v^1, v^2, v^3) = \dot{\mathbf{x}} = (\dot{x}^1, \dot{x}^2, \dot{x}^3)$. Dot denotes $\partial/\partial t$.

We have a domain C in \mathbf{u} -space, and for each t an imbedding $\phi_t: C \rightarrow D_t$ of C into x -space. The mapping $(\mathbf{u}, t) \rightarrow \phi_t(\mathbf{u})$ is assumed twice continuously differentiable, and we write $\phi_t(\mathbf{u}) = \mathbf{x}(\mathbf{u}, t)$, the Lagrange description.

For fixed t , the Jacobian matrix of ϕ_t will be written

$$\frac{\partial \mathbf{x}}{\partial \mathbf{u}} = \left[\frac{\partial x^i}{\partial u^j} \right].$$

It is non-singular everywhere, and its inverse is $\partial \mathbf{u} / \partial \mathbf{x} = [\partial u^j / \partial x^i]$. Its determinant $|\partial \mathbf{x} / \partial \mathbf{u}|$ is usually called the Jacobian of ϕ_t .

We shall need a useful formula from determinant theory. If $A = A(t)$ is a non-singular matrix function, then

$$(4.2) \quad \frac{|A|}{|A|} = \text{trace}(AA^{-1}).$$

We apply (4.2) to the Jacobian matrix. First we note that $(\partial x^i / \partial u^j) \cdot \partial x^i / \partial u^j = \partial v^i / \partial u^j$, hence

$$\begin{aligned} \text{trace} \left\{ \left(\frac{\partial \mathbf{x}}{\partial \mathbf{u}} \right) \cdot \left(\frac{\partial \mathbf{x}}{\partial \mathbf{u}} \right)^{-1} \right\} &= \text{trace} \left\{ \left[\frac{\partial v^i}{\partial u^j} \right] \left[\frac{\partial u^j}{\partial x^k} \right] \right\} \\ &= \sum_{i,j} \frac{\partial v^i}{\partial u^j} \frac{\partial u^j}{\partial x^i} = \sum \frac{\partial v^i}{\partial x^i} = \text{div } \mathbf{v}. \end{aligned}$$

The result is

$$(4.3) \quad \frac{d}{dt} \left| \frac{\partial \mathbf{x}}{\partial \mathbf{u}} \right| = \left| \frac{\partial \mathbf{x}}{\partial \mathbf{u}} \right| (\text{div } \mathbf{v}).$$

Now set

$$f(t) = \iiint_{D_t} F(\mathbf{x}, t) dx^1 dx^2 dx^3.$$

By the change of variables rule,

$$f(t) = \iiint_C F[\mathbf{x}(\mathbf{u}, t), t] \left| \frac{\partial \mathbf{x}}{\partial \mathbf{u}} \right| du^1 du^2 du^3.$$

Differentiation of this fixed domain integral is routine. We use (4.3) and then change back to D_t as soon as possible:

$$\begin{aligned} \dot{f}(t) &= \iiint_C \left\{ \left[\sum \frac{\partial F}{\partial x^i} v^i + \frac{\partial F}{\partial t} \right] \left| \frac{\partial \mathbf{x}}{\partial \mathbf{u}} \right| + F[\mathbf{x}, t] \left| \frac{\partial \mathbf{x}}{\partial \mathbf{u}} \right| (\text{div } \mathbf{v}) \right\} du^1 du^2 du^3 \\ &= \iiint_{D_t} \left\{ (\text{grad } F) \cdot \mathbf{v} + F \text{div } \mathbf{v} + \frac{\partial F}{\partial t} \right\} dx^1 dx^2 dx^3. \end{aligned}$$

But $(\text{grad } F) \cdot \mathbf{v} + F \text{ div } \mathbf{v} = \text{div } (F\mathbf{v})$, so formula (4.1) follows. The proof is not overwhelming once the ground has been paved.

5. Flux across a moving surface. Suppose in a region of \mathbf{x} -space we have a piece of surface S_t that moves with time. We assume that S_t is oriented, with vectorial area element $d\boldsymbol{\sigma}$, and that S_t is described by a map $(\mathbf{u}, t) \rightarrow \mathbf{x}(\mathbf{u}, t)$, where $\mathbf{u} = (u^1, u^2)$ varies over a domain C in the \mathbf{u} -plane. The surface might also be considered as moving with a flow velocity $\mathbf{v} = \mathbf{v}(\mathbf{x}, t)$ as in Section 4. Suppose $\mathbf{F}(\mathbf{x}, t)$ is a time dependent vector field in the region, and set

$$f(t) = \iint_{S_t} \mathbf{F} \cdot d\boldsymbol{\sigma},$$

so that $f(t)$ is the flux of the vector field \mathbf{F} across the moving surface. The problem is to find $\dot{f}(t)$. Now obviously this is fresh ground. First of all, the domain of integration has smaller dimension than does the ambient space. Next, if we take the physicist's point of view, and compare S_t with S_{t+dt} (as in Fig. 2), there won't be an overlap in general, so we must expect a more complicated differentiation formula. In fact, the formula is

$$(5.1) \quad \frac{d}{dt} \iint_{S_t} \mathbf{F} \cdot d\boldsymbol{\sigma} = \iint_{S_t} (\text{div } \mathbf{F}) \mathbf{v} \cdot d\boldsymbol{\sigma} - \int_{\partial S_t} (\mathbf{v} \times \mathbf{F}) \cdot d\mathbf{x} + \iint_{S_t} \mathbf{F} \cdot d\boldsymbol{\sigma}.$$

We might have guessed the second and third terms on the right because of (3.3), but the first term could not have been predicted from the previous discussion. Formula (5.1), with a physicist's proof, appears in Abraham and Becker [1]. The method used to prove (4.1) is inadequate for proving (5.1). It is interesting to try it (formally) because it leads to the wrong answer and provides a good lesson in the care that must be exercised with several variable transformations.

Instead of proving (5.1), we shall pass on to its natural generalization, concerned with a moving r -domain in n -space.

6. Interior product. More than half the job of proving a generalization of (5.1) is formulating the result in a tractable language. First we must drop the idea of integrating a *function* with respect to a *measure*. What we integrate is an exterior differential form over an oriented field of integration (oriented chain). (Particular care must be taken with orientation, because it is so easy to get incorrect signs.) As soon as we take this new point of view, we see that the result we are after has nothing to do with the euclidean structure of space. The result is meaningful for any coordinate space, more generally for a differentiable manifold with no additional structure whatever. For an exposition of the theory of differential forms and their integrals, see any modern book on differential geometry or advanced calculus, especially Flanders [2].

A reasonable formulation of (5.1) in higher space necessarily uses some notation

and some operations. One operation that is not widely known is the interior product, whereby a vector field and a p -form contract to a $(p-1)$ -form.

If \mathbf{v} is a vector field and α is a one-form, we write the effect of α on \mathbf{v} (the dual pairing) as $\langle \mathbf{v}, \alpha \rangle$. Thus

$$\langle \sum v^i \frac{\partial}{\partial x^i}, \sum a_j dx^j \rangle = \sum v^i a_i.$$

The **interior product** of \mathbf{v} and a decomposable p -form $\omega = \alpha^1 \wedge \alpha^2 \wedge \cdots \wedge \alpha^p$ is defined by

$$(6.1) \quad \mathbf{v} \lrcorner (\alpha^1 \wedge \cdots \wedge \alpha^p) = \sum (-1)^{i-1} \langle \mathbf{v}, \alpha^i \rangle \alpha^1 \wedge \cdots \wedge \alpha^{i-1} \wedge \alpha^{i+1} \wedge \cdots \wedge \alpha^p.$$

By linearity, $\mathbf{v} \lrcorner \omega$ is extended to all p -forms ω . To prove that (6.1) really defines an operation that is independent of the representation of ω as a linear combination of decomposable p -forms, it suffices to observe that the right-hand side of (6.1) is an alternating multilinear function of $(\alpha^1, \dots, \alpha^p)$.

Here are some examples. We set

$$\mathbf{v} = u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + w \frac{\partial}{\partial z}.$$

(To free ourselves of the euclidean "length and direction" concept of a vector, we consider a vector as a directional differentiation.) Then

$$(6.2) \quad \begin{cases} \mathbf{v} \lrcorner (F dx + G dy + H dz) = uF + vG + wH, \\ \mathbf{v} \lrcorner (F dy \wedge dz + G dz \wedge dx + H dx \wedge dy) \\ \quad = (wG - vH) dx + (uH - wF) dy + (vF - uG) dz, \\ \mathbf{v} \lrcorner (F dx \wedge dy \wedge dz) = F(u dy \wedge dz + v dz \wedge dx + w dx \wedge dy). \end{cases}$$

We may express these formulas in ordinary vector notation. Set $\mathbf{F} = (F, G, H)$. Then

$$(6.3) \quad \begin{cases} \mathbf{v} \lrcorner (\mathbf{F} \cdot d\mathbf{x}) = \mathbf{v} \cdot \mathbf{F} \\ \mathbf{v} \lrcorner (\mathbf{F} \cdot d\boldsymbol{\sigma}) = -(\mathbf{v} \times \mathbf{F}) \cdot d\mathbf{x} \\ \mathbf{v} \lrcorner (F dx \wedge dy \wedge dz) = F \mathbf{v} \cdot d\boldsymbol{\sigma}. \end{cases}$$

We mention in passing two easily proved formulas:

$$\begin{aligned} \mathbf{v} \lrcorner (\omega \wedge \eta) &= (\mathbf{v} \lrcorner \omega) \wedge \eta + (-1)^{\deg \omega} \omega \wedge (\mathbf{v} \lrcorner \eta), \\ \mathbf{u} \lrcorner (\mathbf{v} \lrcorner \omega) &= -\mathbf{v} \lrcorner (\mathbf{u} \lrcorner \omega). \end{aligned}$$

7. The general Leibniz rule. We are given a p -dimensional time-dependent chain (field of integration) D_t in n -space. We think of D_t as a given by a map

$$\phi: (\mathbf{u}, t) \rightarrow \mathbf{x}(\mathbf{u}, t),$$

where \mathbf{u} runs over a fixed domain C in the p -dimensional \mathbf{u} -space.

We also have an exterior p -form ω whose coefficients are time-dependent. In local coordinates,

$$(7.1) \quad \omega = \sum a_H(x, t) dx^H, \quad dx^H = dx^{h_1} \wedge \cdots \wedge dx^{h_p},$$

where $1 \leq h_1 < h_2 < \cdots < h_p \leq n$. We seek the derivative of $\int_{D_t} \omega$. The answer is

$$(7.2) \quad \frac{d}{dt} \int_{D_t} \omega = \int_D \mathbf{v} \lrcorner d_{\mathbf{x}} \omega + \int_{\partial D_t} \mathbf{v} \lrcorner \omega + \int_{D_t} \dot{\omega}.$$

Here $\dot{\omega} = \sum \dot{a}_H dx^H$ if ω is represented by (7.1). The exterior derivative $d_{\mathbf{x}} \omega$ is taken with respect to the space variables only. (Actually it would not matter if we included the dt term in $d\omega$ because $\mathbf{v} \lrcorner$ would wipe it out.) Precisely, $d\omega = d_{\mathbf{x}} \omega + dt \wedge \dot{\omega}$ in (\mathbf{x}, t) -space. As before $\mathbf{v} = \dot{\mathbf{x}}$.

Formula (7.2) has an attractive simplicity, and the presence of an exterior derivative suggests that its proof involves Stokes's theorem. Such a proof is not hard in itself, but requires careful preparation. We note that the other versions of the Leibniz rule we have discussed are all special cases of (7.2). This statement follows readily from (6.3).

Here is yet another special case. Let C_t be a moving curve in 3-space, so $\partial C_t = \{\mathbf{x}_1(t)\} - \{\mathbf{x}_0(t)\}$. The motion is described by a velocity vector

$$\mathbf{v} = u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + w \frac{\partial}{\partial z},$$

and $\mathbf{v}[\mathbf{x}_0(t), t] = \dot{\mathbf{x}}_0$, $\mathbf{v}[\mathbf{x}_1(t), t] = \dot{\mathbf{x}}_1$. We want

$$\frac{d}{dt} \int_{C_t} \omega, \quad \text{where } \omega = \mathbf{F} \cdot d\mathbf{x}.$$

In the case of this line integral, $d_{\mathbf{x}} \omega = (\text{curl } \mathbf{F}) \cdot d\boldsymbol{\sigma}$, and by (6.3),

$$\mathbf{v} \lrcorner \omega = \mathbf{v} \cdot \mathbf{F}, \quad \mathbf{v} \lrcorner d_{\mathbf{x}} \omega = -[\mathbf{v} \times (\text{curl } \mathbf{F})] \cdot d\mathbf{x}.$$

Therefore (7.2) specializes to

$$(7.5) \quad \begin{aligned} \frac{d}{dt} \int_{C_t} \mathbf{F} \cdot d\mathbf{x} &= - \int_{C_t} [\mathbf{v} \times (\text{curl } \mathbf{F})] \cdot d\mathbf{x} \\ &+ \left\{ \mathbf{F}[\mathbf{x}_1(t), t] \cdot \dot{\mathbf{x}}_1(t) - \mathbf{F}[\mathbf{x}_0(t), t] \cdot \dot{\mathbf{x}}_0(t) \right\} \\ &+ \int_{C_t} \dot{\mathbf{F}} \cdot d\mathbf{x}. \end{aligned}$$

8. Proof of (7.2). There is a technical advantage in taking the time variables

first: signs are simplified. Thus we have

$$\phi: [a, b] \times C \rightarrow R^n,$$

where $[a, b]$ is a closed interval on the t -axis, and C is a p -chain in R^p , the \mathbf{u} -space. We assume ϕ continuously differentiable, so it is actually defined on an open neighborhood of $[a, b] \times C$. We shall use the boundary formula

$$\begin{aligned} (8.1) \quad \partial([a, b] \times C) &= (\partial[a, b]) \times C - [a, b] \times \partial C \\ &= \{b\} \times C - \{a\} \times C - [a, b] \times \partial C. \end{aligned}$$

We must review the process of integrating an exterior p -form over an (oriented differentiable singular) p -chain. Let α be a p -form in R^n and $\psi: C \rightarrow R^n$ a p -chain into the domain of α . The defining formula for integral is

$$\int_{\psi_*(C)} \alpha = \int_C \psi^*(\alpha),$$

where $\psi^*(\alpha)$ is the p -form on C induced by ψ , so $\psi^*(\alpha) = A(\mathbf{u}) du^1 \wedge \cdots \wedge du^p$. Then

$$\int_C \psi^*(\alpha) = \int_C A(u) du^1 du^2 \cdots du^p$$

is an ordinary (Riemann) integral, and it may be iterated in any order. For example if C is a rectangle,

$$\begin{aligned} \iint_C A(u^1, u^2) du^2 \wedge du^1 &= - \iint_C A du^1 \wedge du^2 = \iint_C A du^1 du^2 \\ &= \int_a^b du^1 \int_c^d A du^2 = \int_c^d du^2 \int_a^b A du^1. \end{aligned}$$

In (7.2), the last term, $\int \dot{\omega}$, results from integrand variation only. As before, we shall use the chain rule for this part of the formula, thereby reducing to the case $\omega = \sum a_H(\mathbf{x}) dx^H$. This saves the introduction of additional spaces and mappings; there will be quite enough as it is.

We write $\mathbf{x} = \mathbf{x}(t, \mathbf{u}) = \phi(t, \mathbf{u})$ and $\mathbf{v} = \dot{\mathbf{x}} = \partial \mathbf{x} / \partial t$. We also introduce

$$\phi_t: C \rightarrow R^n, \quad \phi_t(\mathbf{u}) = \phi(t, \mathbf{u}).$$

Each p -form on C may be considered as a p -form on $[a, b] \times C$ via the projection $(t, \mathbf{u}) \rightarrow \mathbf{u}$. In particular we shall consider $\phi_t^* \omega$ as a p -form on $[a, b] \times C$. We state two essential formulas:

$$\begin{aligned} (8.2) \quad \phi^* \omega &= \phi_t^* \omega + dt \wedge \phi_t^*(\mathbf{v} \lrcorner \omega) \\ d(\phi^* \omega) &= dt \wedge \phi_t^*(\mathbf{v} \lrcorner d\omega). \end{aligned}$$

Their proof is based on the decomposition $\phi^*(d\mathbf{x}) = \phi_t^*(d\mathbf{x}) + \mathbf{v} dt$ of $\phi^*(d\mathbf{x})$ into the

part involving the space variables du^j and part involving dt . Then, for example,

$$\begin{aligned}\phi^*(dx^1 \wedge \cdots \wedge dx^q) &= \phi^*(dx^1) \wedge \cdots \wedge \phi^*(dx^q) \\ &= (\phi_t^* dx^1 + v^1 dt) \wedge \cdots \wedge (\phi_t^* dx^q + v^q dt) \\ &= \phi_t^* dx^1 \wedge \cdots \wedge \phi_t^* dx^q + dt \wedge [v^1 \phi_t^* dx^2 \wedge \cdots \wedge \phi_t^* dx^q \\ &\quad + \cdots + (-1)^{q-1} v^q \phi_t^* dx^1 \wedge \cdots \wedge \phi_t^* dx^{q-1}].\end{aligned}$$

The first formula follows easily. Now apply it to $d\omega$, noting that $\phi_t^*(d\omega)$ is a $(p+1)$ -form on C , hence 0, so that $\phi^*(d\omega) = dt \wedge \phi_t^*(\mathbf{v} \lrcorner d\omega)$. But $d(\phi^*\omega) = \phi^*(d\omega)$.

Now we use Stokes's theorem:

$$(8.3) \quad \int_{[a,t] \times C} d(\phi^*\omega) = \int_{\partial([a,t] \times C)} \omega.$$

On the left,

$$\int_{[a,t] \times C} d(\phi^*\omega) = \int_{[a,t] \times C} ds \wedge \phi_s^*(\mathbf{v} \lrcorner d\omega) = \int_a^t ds \int_C \phi_s^*(\mathbf{v} \lrcorner d\omega).$$

On the right, we have three terms according to (8.1). On the bases $\{a\} \times C$ and $\{t\} \times C$ of the cylinder t is constant, so $dt \wedge (\quad)$ in (8.2) makes no contribution. On the lateral side $[a,t] \times \partial C$ of the cylinder, the p -form $\phi_t^*\omega = 0$, because it is 0 on the $(p-1)$ -chain ∂C . Therefore

$$\begin{aligned}\int_{\partial([a,t] \times C)} \phi^*\omega &= \int_{[t] \times C} \phi_s^*\omega - \int_{[a] \times C} \phi_s^*\omega - \int_{[a,t] \times \partial C} ds \wedge \phi_s^*(v \lrcorner \omega) \\ &= \int_C \phi_t^*\omega - \int_C \phi_a^*\omega - \int_a^t ds \int_{\partial C} \phi_s^*(\mathbf{v} \lrcorner \omega).\end{aligned}$$

Hence (8.3) implies

$$\int_C \phi_t^*\omega - \int_C \phi_a^*\omega = \int_a^t ds \int_C \phi_s^*(\mathbf{v} \lrcorner d\omega) + \int_a^t ds \int_{\partial C} \phi_s^*(\mathbf{v} \lrcorner \omega),$$

that is,

$$(8.5) \quad \int_{D_t} \omega - \int_{D_a} \omega = \int_a^t ds \int_{D_s} \mathbf{v} \lrcorner d\omega + \int_a^t ds \int_{\partial D_s} \mathbf{v} \lrcorner \omega.$$

We summon the fundamental theorem once again:

$$\frac{d}{dt} \int_{D_t} \omega = \int_{D_t} \mathbf{v} \lrcorner d\omega + \int_{\partial D_t} \mathbf{v} \lrcorner \omega.$$

This completes the proof and our story.

REMARK: Because the terms in (7.2) are additive in D_i , the formula is valid for the most general p -chain, a linear combination of coordinatized ones.

References

1. M. Abraham and R. Becker, *Classical Theory of Electricity and Magnetism*, 2nd ed., Blackie, London, 1950, pp. 39–40.
2. H. Flanders, *Differential Forms with Applications to the Physical Sciences*, Academic Press, New York, 1963.
3. W. Fleming, *Functions of Several Variables*, Addison-Wesley, Reading, Mass., 1965, pp. 197–200.
4. L. H. Loomis and S. Sternberg, *Advanced Calculus*, Addison-Wesley, Reading, Mass., 1968, pp. 419 and 456.
5. W. Prager, *Introduction to Mechanics of Continua*, Ginn, Boston, 1961, pp. 75–76.
6. I. S. Sokolnikoff and R. M. Redheffer, *Mathematics of Physics and Engineering*, 2nd ed., McGraw-Hill, New York, 1966, pp. 428–430.

THE STANFORD UNIVERSITY COMPETITIVE EXAMINATION IN MATHEMATICS

G. POLYA, Stanford University, and
J. KILPATRICK, Teachers College, Columbia University

1. Introduction. For twenty years, from 1946 to 1965, the Department of Mathematics at Stanford University conducted a competitive examination for high school seniors. The immediate and principal purpose of the examination was to identify,

Prof. Polya received his Univ. Budapest degree in 1912 and holds honorary degrees from the E. T. H. Zürich, Univ. Alberta, and Univ. Wisconsin. He taught at the E. T. H. until 1940 and has been at Stanford Univ. since. His numerous visiting posts include Cambridge, Oxford, Paris, Göttingen, and Princeton. He is a Correspondent of the Paris Academy of Sciences and holds honorary membership in the Council of the Soc. Math. de France, the London Math. Soc. and the Swiss Math. Soc. Prof. Polya received the M.A.A. Distinguished Service Award in 1963 and the 1968 N. Y. Film Festival top Blue Ribbon for "Let us teach guessing".

The scientific contributions of George Polya include over 230 research papers and the books, *Inequalities* (with Hardy and Littlewood), *How to Solve It*, *Isoperimetric Inequalities* (with Szegő), *Mathematics and Plausible Reasoning* (2 v.), and *Mathematical Discovery* (2v.).

Prof. Polya's personal influence on three generations of mathematicians has been enormous. Perhaps no book in existence has influenced the direction of thinking of young mathematicians more than his two volume masterpiece with G. Szegő, *Aufgaben und Lehrsätze aus der Analysis*.

Jeremy Kilpatrick took the Stanford Examination himself while a senior in high school; later he assisted in grading the Stanford Examination in its last few years. While a graduate student he worked closely with Professor Polya, and he received his Stanford Ph.D. in Education under E. G. Begle. He has since been Assistant and Associate Professor at Teachers College, Columbia. He works in the heuristics of problem solving and in mathematical abilities, and he is the co-editor with I. Wirszup of the series "Soviet Studies in the Psychology of Learning and Teaching Mathematics".
Editor.

among each year's high school graduates, singularly capable students and attract them to Stanford. The broader purpose was to stimulate interest in mathematics among high school students and teachers generally, as well as the public.

The examination was modeled on the Eötvös Competition [see 3], which was organized in Hungary in 1894, and which, in turn, appears to have been suggested by similar competitions in England and France. Gabor Szegő, chairman of the Stanford Department of Mathematics in 1946 and winner of the Eötvös Competition in 1912, initiated the Stanford examination.

The examination was established in the belief that an early manifestation of mathematical ability is a definite indication of exceptional intelligence and suitability for intellectual leadership in any field of endeavor. Furthermore, mathematical ability can be tested at a comparatively early age because it is manifested "not so much by the amount of accumulated knowledge as by the originality of mind displayed in the game of grappling with difficult though elementary problems [2, p. 406]."

As Buck [1] noted some years ago in reviewing mathematical competitions, an examination can be designed, broadly speaking, to test either achievement or aptitude. The Stanford University Competitive Examination in Mathematics was of the latter type. It emphasized

originality and insight rather than routine competence A typical question might call for specific knowledge within the reach of those being tested, but would call for the employment of this in unusual ways requiring a high degree of ingenuity. The question may in fact introduce certain concepts which are quite unfamiliar to the student. In short, the winning student is asked to demonstrate research ability [1, pp. 204–205].

The first Stanford examination, in 1946, was administered in 60 California high schools to 322 participants. The winner was awarded a one-year scholarship by Stanford University; honorable mention and a mathematics book were given to three other participants. In 1953, the examination was extended beyond California to include Arizona, Oregon, and Washington; the number of scholarships was increased to two; and the number of honorable mention awards and books was increased to ten or so. From 1958 to 1962, the examination was co-sponsored by Sylvania Electric Products, Inc. The last examination, in 1965, was administered to about 1200 participants in 151 centers in California, Arizona, Idaho, Montana, Nevada, Oregon, and Washington. Cash prizes of \$500, \$250, and \$250 were awarded to the three winners; honorable mention and a mathematics book went to eighteen participants. The examination was discontinued after 1965 mainly because the Stanford Department of Mathematics turned its interest to more graduate teaching.

Announcements of the examination were sent each year to all public and private high schools in each state where the examination was to be administered. Larger schools were designated as centers; students from other schools were free to arrange to take the examination in a convenient location. The examination was administered by teachers and school personnel on a Saturday afternoon in March or April. The

participants were given three hours to attempt three to five problems. The following instructions were given:

No books or notebooks may be used. You may not be able to do all the problems in three hours, but whatever you do should be carefully thought out. Scratch paper may be used. Either pen or pencil may be used. No questions concerning the test should be asked of the person in charge.

Good presentation *counts*!

It should be clear, concise, complete.

The papers were read in a two-stage process: First, they were read by teams of graduate students in the Department of Mathematics, including, as was sometimes possible, graduate students who were experienced high school teachers. Each team of two students was assigned a problem to read in as many papers as they could handle. Papers containing either a stated minimum of good solutions (for example, one and a half or two out of four) or some special feature were forwarded to the second stage. In the second stage, each paper that survived the first screening was read by at least one faculty member of the Department. The papers considered most likely to be winners were read by all participating faculty members.

To make the selection of winners easier, the problems were devised so that only a very few participants would be able to solve all of them. On the other hand, to avoid too much frustration, the first problem was usually more accessible than the others, especially in the later years, so that many participants were able to solve it.

Although the mathematical content of the problems did not go beyond that of the high school curriculum, the problems were of types seldom found in textbooks. The purpose of such problems was not only to test the students' originality, but also to enrich the high school mathematics program by suggesting some new directions for students' and teachers' work. The types of problems included (1) "guess and prove," in which one first guesses and then proves a mathematical fact, (2) "test consequences," in which one tests the consequences of a general statement, (3) "you may guess wrong," in which a highly plausible guess is incorrect, (4) "small scale theory," in which a sequence of subproblems illustrates theory construction, and (5) "red herring," in which an obvious relationship among the data turns out to be irrelevant to the solution [see 6, pp. 160–161, ex. 1; 8, p. 139, ex. 14.23].

The problems were of the sort used as illustrations in *How to Solve It* [4], the two volumes of *Mathematics and Plausible Reasoning* [5, 6], and the two volumes of *Mathematical Discovery* [7, 8]. In fact, many of the problems appear, usually with solutions, in these books. The interested reader is directed, in particular, to the following sources:

1. Part IV of [4] contains problems taken (with a few minor changes) from the 1946–56 examinations. Hints and solutions are provided.

2. Chapters I and VII of [5] may be useful in attacking problems involving induction. Several of the problems from the 1946–50 examinations are included among the examples and comments at the end of the chapters.

3. The examples and comments on Chapter XVI of [6] contain some problems from the 1946–52 examinations. The appendix added in the second edition contains additional problems, one of which is taken from the 1958 and one from the 1964 examination.

4. Chapters 2 and 6 of [7] discuss and elaborate Descartes' method for solving problems. Some of the examples used to illustrate the method, and several of the problems at the end of the chapters, are taken from the 1951–61 examinations. The appendix of [7] gives some suggestions for teachers on how to use such problems in class.

5. Chapter 15 of [8] illustrates the use of research problems—including one from the 1965 examination—to provide students something of an opportunity for independent creative work. Additional problems from the 1961–65 examinations are given at the end of the chapter and in the appendix to the corrected printing.

Each year the examination was conducted (with six exceptions), an article listing the problems and the winners was published either in this MONTHLY or in the *California Mathematics Council Bulletin*. Some hints and solutions also were published in the latter journal. The problems have never before, however, been collected together as a set.

Section 2 contains the complete set of problems from the Stanford University Competitive Examination in Mathematics. They are numbered as follows: “46.1” refers to problem 1 in the 1946 examination. Owing to space limitations, hints and solutions are not provided. A booklet containing the problems and a complete set of hints and solutions, many developed in seminars on problem solving at Stanford University and at Teachers College, together with specific references to articles and books in which the problems have previously appeared, will be published under the title *The Stanford Mathematics Problem Book* by the Teachers College Press, Teachers College, Columbia University, New York, NY 10027.

2. Problems

46.1. In a tennis tournament there are $2n$ participants. In the first round of the tournament each participant plays just once, so there are n games, each occupying a pair of players. Show that the pairing for the first round can be arranged in exactly

$$1 \times 3 \times 5 \times 7 \times 9 \cdots \times (2n - 1)$$

different ways.

46.2. In a tetrahedron (which is not necessarily regular) two opposite edges have the same length a and they are perpendicular to each other. Moreover they are each perpendicular to a line of length b which joins their midpoints. Express the volume of the tetrahedron in terms of a and b , and prove your answer.

46.3. Consider the following four propositions, which are not necessarily true.

I. If a polygon inscribed in a circle is equilateral it is also equiangular.

II. If a polygon inscribed in a circle is equiangular it is also equilateral.

III. If a polygon circumscribed about a circle is equilateral it is also equiangular.

IV. If a polygon circumscribed about a circle is equiangular it is also equilateral.

(A) State which of the four propositions are true and which are false, giving a proof of your statement in each case.

(B) If, instead of general polygons, we should consider only quadrilaterals which of the four propositions are true and which are false? And if we consider only pentagons?

In answering (B) you may state conjectures, but prove as much as you can and separate clearly what is proved and what is not.

47.1. To number the pages of a bulky volume the printer used 1890 digits. How many pages has the volume?

47.2. Among grandfather's papers a bill was found:

72 turkeys \$-67.9-

The first and last digit of the number that obviously represented the total price of those fowls are replaced here by blanks, for they have faded and are now illegible.

What are the two faded digits and what was the price of one turkey?

47.3. Determine m so that the equation in x

$$x^4 - (3m + 2)x^2 + m^2 = 0$$

has four real roots in arithmetic progression.

47.4. Let α , β and γ denote the angles of a triangle. Show that

$$\sin \alpha + \sin \beta + \sin \gamma = 4 \cos \frac{\alpha}{2} \cos \frac{\beta}{2} \cos \frac{\gamma}{2},$$

$$\sin 2\alpha + \sin 2\beta + \sin 2\gamma = 4 \sin \alpha \sin \beta \sin \gamma,$$

and

$$\sin 4\alpha + \sin 4\beta + \sin 4\gamma = -4 \sin 2\alpha \sin 2\beta \sin 2\gamma.$$

48.1. Consider the table:

$$\begin{array}{rcl} 1 & = & 1 \\ 2 + 3 + 4 & = & 1 + 8 \\ 5 + 6 + 7 + 8 + 9 & = & 8 + 27 \\ 10 + 11 + 12 + 13 + 14 + 15 + 16 & = & 27 + 64 \end{array}$$

Guess the general law suggested by these examples, express it in suitable mathematical notation, and prove it.

48.2. Three numbers are in arithmetic progression, three other numbers in geometric progression. Adding the corresponding terms of these two progressions successively, we obtain

85, 76, and 84

respectively, and adding all three terms of the arithmetic progression, we obtain 126. Find the terms of both progressions.

48.3. From the peak of a mountain, you see two points, A and B , in the plain. The lines of vision, directed to these points, include the angle γ . The inclination of the first line of vision to a horizontal plane is α , that of the second line β . It is known that the points A and B are on the same level and that the distance between them is c .

Express the elevation x of the peak above the common level of A and B in terms of the angles α , β , γ and the distance c .

48.4. A first sphere has the radius r_1 . About this sphere circumscribe a regular tetrahedron. About this tetrahedron circumscribe a second sphere with radius r_2 . About this second sphere circumscribe a cube. About this cube circumscribe a third sphere with radius r_3 .

Find the ratios $r_1 : r_2 : r_3$ (which should be, according to Kepler, the ratios of the mean distances of the planets Mars, Jupiter, and Saturn from the Sun, but which are, in fact, rather different from the true ratios).

49.1. Prove that no number in the sequence

$$11, 111, 1111, 11111, \dots$$

is the square of an integer.

49.2. The three sides of a triangle are of lengths l , m , and n , respectively. The numbers l , m , and n are positive integers,

$$l \leq m \leq n.$$

(A) Take $n = 9$ and find the number of different triangles of the described kind.

(B) Take various values of n and find a general law.

49.3. (A) Prove the following theorem: *A point lies inside an equilateral triangle and has the distances x , y , and z from the three sides respectively; h is the altitude of the triangle. Then $x + y + z = h$.*

(B) State precisely and prove the analogous theorem in solid geometry concerning the distances of an inner point from the four faces of a regular tetrahedron.

(C) Generalize both theorems so that they should apply to any point in the plane or space, respectively (and not only to points inside the triangle or tetrahedron). Give precise statements and, if you have time, also proofs.

50.1. Observe that

$$1 = 1$$

$$1 - 4 = -(1 + 2)$$

$$1 - 4 + 9 = 1 + 2 + 3$$

$$1 - 4 + 9 - 16 = -(1 + 2 + 3 + 4)$$

Guess the general law suggested by these examples, express it in suitable mathematical notation, and prove it.

50.2. Given a square. Find the locus of the points from which the square is seen under an angle (A) of 90° (B) of 45° . (Let P be a point outside the square, but in the same plane. The smallest angle

with vertex P containing the square is the "angle under which the square is seen" from P .) Sketch clearly both loci, give a full description, and a proof.

50.3. Call "axis" of a solid a straight line joining two points of the surface of the solid and such that the solid, rotated about this line through an angle which is greater than 0° and less than 360° coincides with itself.

A cube has 13 different axes, which are of three different kinds. Describe clearly the location of these axes, find the angle of rotation associated with each. Assuming that the edge of the cube is of unit length, compute the arithmetic mean of the lengths of the 13 axes. Do not use tables and compute to two decimals.

51.1. The length of the perimeter of a right triangle is 60 inches and the length of the altitude perpendicular to the hypotenuse is 12 inches. Find the sides of the triangle.

51.2. A quadrilateral is cut into four triangles by its two diagonals. We call two of these triangles "opposite" if they have a common vertex but no common side. Prove the following statements: (A) The product of the areas of two opposite triangles is equal to the product of the areas of the other two opposite triangles. (B) The quadrilateral is a trapezoid if, and only if, there are two opposite triangles equal in area. (C) The quadrilateral is a parallelogram if, and only if, all four triangles are equal in area.

51.3. We consider the frustum of a right circular cone. The plane that is parallel to the lower and upper bases of the frustum and at equal distance from both intersects the frustum in the "median circle." The frustum and a cylinder have the same altitude, and the median circle of the frustum is the base of the cylinder. Which one of these two solids has the greater volume, the frustum or the cylinder? Prove your answer! (A possible proof is by algebra: Express both volumes in terms of suitable data and transform their difference so that its sign becomes obvious.)

52.1. Prove the proposition: If a side of a triangle is less than the average (arithmetic mean) of the two other sides, the opposite angle is less than the average of the two other angles.

52.2. Consider the frustum of a right pyramid with square base. Call "midsection" the intersection of the frustum with a plane parallel to the base and the top and at the same distance from both. Call "intermediate rectangle" the rectangle of which one side is equal to a side of the base and the other side is equal to a side of the top.

Four different friends of yours agree that the volume of the frustum equals the altitude multiplied by a certain area, but they disagree and make four different proposals regarding this area:

- I. the midsection,
- II. the average of the base and the top,
- III. the average of the base, the top, and the midsection,
- IV. the average of the base, the top, and the intermediate rectangle,

Let h be the altitude of the frustum, a the side of its base, and b the side of its top. Express each of the four proposed rules in mathematical notation, decide whether it is right or wrong, and prove your answer.

52.3. Prove that the only solution of the equation

$$x^2 + y^2 + z^2 = 2xyz$$

in integers x , y , and z is $x = y = z = 0$.

53.1. Bob has 10 pockets and 44 silver dollars. He wants to put his dollars into his pockets so distributed that each pocket contains a different number of dollars.

(A) Can he do so?

(B) Generalize the problem, considering p pockets and n dollars. The problem is the most interesting when

$$n = \frac{(p+1)(p-2)}{2}.$$

Why?

53.2. Observe that the value of

$$\frac{1}{2!} + \frac{2}{3!} + \frac{3}{4!} + \cdots + \frac{n}{(n+1)!}$$

is $1/2$, $5/6$, $23/24$ for $n = 1, 2, 3$, respectively, guess the general law (by observing more values if necessary) and prove your guess.

53.3. Find x , y , u , and v satisfying the system of four equations

$$x + 7y + 3v + 5u = 16$$

$$8x + 4y + 6v + 2u = -16$$

$$2x + 6y + 4v + 8u = 16$$

$$5x + 3y + 7v + u = -16$$

(This may look long and boring: look for a shortcut.)

53.4. The four points G , H , V , and U are (in this order) the four corners of a quadrilateral. A surveyor wants to find the length $UV = x$. He knows the length $GH = l$ and measures the four angles

$$\angle GUH = \alpha, \angle HUV = \beta, \angle UVG = \gamma, \angle GVH = \delta.$$

(A) Express x in terms of α , β , γ , δ , and l .

(B) Find some way to test the correctness of the result.

(C) If you had a clear plan to do (A) characterize it in one short sentence.

54.1. Consider the table

$$\begin{array}{rcl} 1 & = & 1 \\ 3 + 5 & = & 8 \\ 7 + 9 + 11 & = & 27 \\ 13 + 15 + 17 + 19 & = & 64 \\ 21 + 23 + 25 + 27 + 29 & = & 125 \end{array}$$

Guess the general law suggested by these examples, express it in suitable mathematical notation, and prove it.

54.2. The side of a regular hexagon is of length n (n is an integer). By equidistant parallels to its sides, the hexagon is divided into T equilateral triangles each of which has sides of length 1. Let V

denote the number of vertices appearing in this division, and L the number of boundary lines of length 1. (A boundary line belongs to one or two triangles, a vertex to two or more triangles.) When $n = 1$, which is the simplest case, $T = 6$, $V = 7$, $L = 12$.

Consider the general case and express T , V , and L in terms of n . (Guessing is good, proving is better.)

54.3. Show that it is impossible to find (real or complex) numbers a, b, c, A, B , and C such that the equation

$$x^2 + y^2 + z^2 = (ax + by + cz)(Ax + By + Cz)$$

holds identically for independently variable x, y , and z .

55.1. Bob wants a piece of land, exactly level, which has four boundary lines. Two boundary lines run exactly north-south, the two others exactly east-west, and each boundary line measures exactly 100 feet. Can Bob buy such a piece of land in the U.S.? State your reasons!

55.2. (A) Find three numbers p, q , and r so that the equation

$$x^4 + 4x^3 - 2x^2 - 12x + 9 = (px^2 + qx + r)^2$$

holds identically for variable x .

(B) This problem requires the "exact" extraction of a square root of a given polynomial of degree 4, which may be possible in the present case, yet usually it is not. Why not?

55.3. Bob, Peter, and Paul travel together. Peter and Paul are good hikers; each walks p miles per hour. Bob has a bad foot and drives a small car in which two people can ride, but not three; the car covers c miles per hour. The three friends adopted the following scheme: They start together, Paul rides in the car with Bob, Peter walks. After a while, Bob drops Paul who walks on; Bob returns to pick up Peter, and then Bob and Peter ride in the car till they overtake Paul. At this point, they change: Paul rides and Peter walks just as they started and the whole procedure is repeated as often as necessary.

(A) How much progress (how many miles) does the company make per hour?

(B) Through which fraction of the travel time does the car carry just one man?

(C) Check the extreme cases $p = 0$ and $p = c$.

55.4. The vertex of a pyramid opposite the base is called the *apex*. (A) Let us call a pyramid "isosceles" if its apex is at the same distance from all *vertices* of the base. Adopting this definition, prove that the base of an isosceles pyramid is *inscribed* in a circle the center of which is the foot of the pyramid's altitude.

(B) Now let us call a pyramid "isosceles" if its apex is at the same (perpendicular) distance from all *sides* of the base. Adopting this definition (different from the foregoing) prove that the base of an isosceles pyramid is *circumscribed* about a circle the center of which is the foot of the pyramid's altitude.

56.1. Given a regular hexagon and a point in its plane. Draw a straight line through the given point that divides the given hexagon into two parts of equal area.

56.2. I say that you can pay 50 cents in exactly 50 different manners. (The "manner" depends on how many coins of each kind — cents, nickels, dimes, quarters, half dollars — you use.) In how many manners can you pay 25 cents? Am I right about 50 cents? Justify your answer as clearly as you can.

56.3. Construct a hexagon by adding to an arbitrarily given triangle Δ three exterior isosceles triangles each of which has an angle of 120° opposite to that side of Δ that forms its base. Show that

those three vertices of the hexagon that are not vertices of the given Δ are the vertices of an *equilateral* triangle. (It is enough to express just one side s of the allegedly equilateral triangle in terms of the sides a , b , and c of Δ , provided that this expression for s is symmetric in a , b , and c .)

56.4. Ten people are sitting around a round table. The sum of ten dollars is to be distributed among them according to the rule that each person receives one half of the sum that his two neighbors receive jointly. Is there just one way to distribute the money? Prove your answer.

57.1. Bob's stamp collection consists of three books. Two tenths of his stamps are in the first book, several sevenths in the second book, and there are 303 stamps in the third book. How many stamps has Bob? (Is the condition sufficient to determine the unknown?)

57.2. We call a vertex of a tetrahedron *trirectangular* if the three edges starting from it are perpendicular to each other. Given the areas A , B , and C of the three faces adjacent to the trirectangular vertex of a tetrahedron, find the area D of the fourth face, opposite to that vertex. (Which problem of plane geometry would you regard as analogous?)

57.3. Divide a given triangle by three straight cuts into seven pieces four of which are triangles (and the remaining three pentagons). One of the triangular pieces is included by the three cuts, each of the three other triangular pieces is included by a certain side of the given triangle and two cuts.

(A) Choose the three cuts so that the four triangular pieces turn out to be congruent. Describe your choice precisely and draw a clear figure.

(B) Which fraction of the area of the given triangle is the area of a triangular piece in the dissection that you chose?

(It may be advantageous to examine first a particular shape of the given triangle for which the solution is particularly easy.)

58.1. How old is the captain, how many children has he, and how long is his boat? Given the product 32118 of the three desired numbers (integers). The length of the boat is given in feet (is several feet), the captain has both sons and daughters, he has more years than children, but he is not yet one hundred years old. (Give reasons for your answer.)

58.2. Find x , y , u , and v satisfying the system of four equations:

$$x + y + u = 4$$

$$y + u + v = -5$$

$$u + v + x = 0$$

$$v + x + y = -8$$

(This may look long and boring: look for a shortcut.)

58.3. "In any triangle the sum of the three... is greater than the semiperimeter."

Replace the dots ... successively by

I. altitudes

II. medians

III. bisectors (of the angles).

You obtain so three different assertions. Examine each assertion: is it true or false? Prove your answer!

58.4. Observe that the value of

$$1!1 + 2!2 + 3!3 + \cdots + n!n$$

is 1, 5, 23, 119 for $n = 1, 2, 3, 4$, respectively. Guess the general law (by observing more values if necessary) and prove your guess.

59.1. Al and Bill live at opposite ends of the same street. Al had to deliver a parcel at Bill's home, Bill one at Al's home. They started at the same moment, each walked at constant speed and returned home immediately after leaving the parcel at its destination. They met the first time at the distance of a yards from Al's home and the second time at the distance of b yards from Bill's home.

(A) How long is the street?

(B) If $a = 300$ and $b = 400$ who walks faster?

59.2. Pennies (equal circles) are arranged in a regular pattern all over a very-very large table (the infinite plane). We examine two patterns.

In the first pattern, each penny touches four other pennies and the straight lines joining the centers of the pennies in contact dissect the plane into equal squares.

In the second pattern, each penny touches six other pennies and the straight lines joining the centers of the pennies in contact dissect the plane into equal equilateral triangles.

Compute the percentage of the plane covered by pennies (circles) for each pattern.

59.3. Prove: If n is an integer greater than 1, $n^{n-1} - 1$ is divisible by $(n - 1)^2$.

59.4. Erect an (exterior) square on each side of an (arbitrarily given) triangle. Those 6 vertices of these 3 squares that do not coincide with a vertex of the triangle form a hexagon. Three sides of this hexagon are, of course, equal to the corresponding sides of the triangle. Show that each one of the remaining three sides equals the double of a median of the triangle.

60.1. A certain make of ball point pen was priced 50 cents in the store opposite the high school but found few buyers. When, however, the store had reduced the price, the whole remaining stock was sold for \$31.93. What was the reduced price? (Is the condition sufficient to determine the unknown?)

60.2. The point P is so located in the interior of a rectangle that the distance of P from a corner of the rectangle is 5 yards, from the opposite corner 14 yards, and from a third corner 10 yards. What is the distance of P from the fourth corner?

60.3. Prove the identity

$$\cos \frac{a}{2} \cos \frac{a}{4} \cos \frac{a}{8} = \frac{\sin a}{8 \sin a/8}$$

and generalize.

60.4. Of twelve congruent equilateral triangles eight are the faces of a regular octahedron and four the faces of a regular tetrahedron. Find the ratio of the volume of the octahedron to the volume of the tetrahedron.

61.1. Solve the following system of three equations for the unknowns x , y , and z :

$$5732x + 2134y + 2134z = 7866,$$

$$2134x + 5732y + 2134z = 670,$$

$$2134x + 2134y + 5732z = 11464.$$

61.2. It was a very hot day and the 4 couples drank together 44 bottles of coca-cola. Ann had 2,

Betty 3, Carol 4 and Dorothy 5 bottles. Mr. Brown drank just as many bottles as his wife, but each of the other men drank more than his wife: Mr. Green twice, Mr. White three times and Mr. Smith four times as many bottles. Tell the last names of the four ladies. (Prove your answer.)

61.3. Solve the following system of three equations for the unknowns x , y and z (a , b and c are given):

$$x^2y^2 + x^2z^2 = axyz,$$

$$y^2z^2 + y^2x^2 = bxyz,$$

$$z^2x^2 + z^2y^2 = cxyz.$$

61.4. A pyramid is called “regular” if its base is a regular polygon and the foot of its altitude is the center of its base. A regular pyramid has a hexagonal base the area of which is one quarter of the total surface-area S of the pyramid. The altitude of the pyramid is h . Express S in terms of h .

62.1. Solve the system

$$2x^2 - 4xy + 3y^2 = 36$$

$$3x^2 - 4xy + 2y^2 = 36$$

(One solution is easy to guess, but you are required to find *all* solutions. Knowledge of analytic geometry is not needed to solve this problem, but may help to understand the result — how?)

62.2. Each of the four numbers, a , b , c , and d , is positive and less than one. Show that not all four products

$$4a(1 - b), 4b(1 - c), 4c(1 - d), 4d(1 - a)$$

are greater than one.

62.3. On each side of a right triangle, erect an exterior square (as it is usually done to illustrate Pythagoras’ theorem). Join the vertex of the triangle’s right angle to the center of the square on the hypotenuse, and join the centers of the squares on the other two sides. Show that the two line segments so obtained are

- (A) perpendicular to each other and
- (B) of equal length.

62.4. Five edges of a tetrahedron are of the same length a , and the sixth edge is of the length b .

- (A) Express the radius of the sphere circumscribed about the tetrahedron in terms of a and b .
- (B) How would you use the result (A) to determine practically the radius of a spherical surface (of a lens)?

63.1. In a right triangle, c is the length of the hypotenuse, a and b are the lengths of the two other sides, and d is the length of the diameter of the inscribed circle. Prove that

$$a + b = c + d.$$

63.2. Show that the expression

$$n^2 (n^2 - 1) (n^2 - 4)$$

is divisible by 360 for $n = 1, 2, 3, \dots$.

63.3. Solve the system of three equations for the unknowns x , y , and z , giving all solutions:

$$x^2 + 5y^2 + 6z^2 + 8(yz + zx + xy) = 36,$$

$$6x^2 + y^2 + 5z^2 + 8(yz + zx + xy) = 36,$$

$$5x^2 + 6y^2 + z^2 + 8(yz + zx + xy) = 36.$$

(One solution is easy to find.)

63.4. The base of a right prism is a regular hexagon, and the height of the prism is equal to the diameter of the circle inscribed in the base. The volume of the prism is equal to the volume of a regular octahedron. Find the ratio of the surface-areas of these two solids.

Observe that the two solids have the same number of faces, and one of them is a regular solid, but the other is not. Any remark?

64.1. A cake has the shape of a right prism with a square base; it has icing on the top as well as on the sides (that is, on the four lateral faces). The altitude of the prism is $5/16$ of the side of its base. Cut the cake into 9 pieces so that each piece has the same amount of cake and the same amount of icing. One of the 9 pieces should be a right prism with a square base with icing only on the top: Compute the ratio of its altitude to a side of its base and give a clear description, with an acceptable sketch, of all 9 pieces.

64.2. Show that each number of the sequence

$$49, 4489, 444889, 44448889, \dots$$

is a perfect square.

64.3. If the area of a triangle is rational (that is, measured by a rational number) there are four thinkable cases: The triangle may have three or two rational sides, or just one or no rational side. Show by (preferably simple) examples that all four cases are actually possible.

64.4. An examination in three subjects, Algebra, Biology, and Chemistry was taken by 41 students. The following table shows how many students failed in each single subject and in their various combinations:

in	A	B	C	AB	AC	BC	ABC
failed	12	5	8	2	6	3	1

(For instance, 5 students failed in Biology, among whom there were 3 failing both in Biology and in Chemistry, and just one of these 3 failed in all three subjects.) How many students passed in all three subjects?

(Can you think of a suitable diagram that would clarify the underlying idea?)

64.5. Let a , b , and c denote the lengths of the sides of a triangle, and d the length of the bisector of the angle opposite to the side of length c , terminated on the side.

(A) Express d in terms of a , b , and c .

(B) Check the expression obtained in as many ways as you can (by particular cases, limiting cases, and so on).

65.1. "How many children have you, and how old are they?" asked the guest, a mathematics teacher.

"I have three boys," said Mr. Smith. "The product of their ages is 72 and the sum of their ages is the street number."

The guest went to look at the entrance, came back and said: "The problem is indeterminate."

"Yes, that is so," said Mr. Smith, "but I still hope that the oldest boy will some day win the Stanford competition."

Tell the ages of the boys, stating your reasons.

65.2. Of a right triangle, given the length of the hypotenuse c and the area A . On each side of the triangle, describe a square exterior to the triangle and consider the least convex figure containing the three squares (formed by a tight rubber band around them): it is a hexagon (which is irregular, has one side in common with each square, and one of its remaining three sides is obviously of length c).

Find the area of the hexagon.

65.3. Let the numbers x , y and 1 measure the lengths of the three sides of some triangle and suppose that

$$x \leq y \leq 1.$$

Let the point (x, y) with rectangular coordinates x and y represent the triangle on a plane. Describe precisely and sketch clearly the set of those points of the plane that, in the manner explained, represent

- (A) triangles,
- (B) isosceles triangles,
- (C) right triangles,
- (D) acute triangles,
- (E) obtuse triangles.

Locate the representative points of still other noteworthy triangular shapes.

65.4. Find the remainder of the division of the polynomial $x + x^9 + x^{25} + x^{49} + x^{81}$ by the polynomial $x^3 - x$.

References

1. R. Creighton Buck, A look at mathematical competitions, this MONTHLY, 66 (1959) 201–212.
2. Department of Mathematics, Stanford University, The Stanford University Mathematics Examination, this MONTHLY, 53 (1946) 406–409.
3. Hungarian Problem Book I, II, Random House, New York, 1963.
4. G. Polya, How to Solve It, 2nd edition, Doubleday Anchor A 93, 1957, Princeton Univ. Press, 1971.
5. ———, Mathematics and Plausible Reasoning, Vol. 1, Princeton Univ. Press, 1954.
6. ———, Mathematics and Plausible Reasoning, Vol. 2, 2nd edition, Princeton Univ. Press, 1968.
7. ———, Mathematical Discovery, Vol. 1, Wiley, New York, 1962.
8. ———, Mathematical Discovery, Vol. 2, corrected printing, Wiley, New York, 1968.

HOW TO CLASSIFY DIFFERENTIAL POLYNOMIALS

REUBEN HERSH, University of New Mexico

Introduction. What can it mean to say that a polynomial is “elliptic,” “hyperbolic,” “parabolic,” if its degree is not equal to two? The intent of this article is to explain the meaning of these words. In doing so, we shall see how the Fourier transformation provides a bridge between elementary algebra and advanced analysis. Natural questions about differential operators lead to interesting problems in elementary algebra. Nothing we discuss in this article is new. Our purpose is to expound some interesting elementary notions which are usually to be found embedded in highly technical treatises.

Consider a polynomial in $n + 1$ variables,

$$(1) \quad Q(a, b) = \sum_{i=1}^l \sum_{j=1}^m c_{ij} a^i b^j,$$

where j is a multi-index $j = j_1, \dots, j_n$, b is an n -tuple, and $b^j = \prod b_i^{j_i}$.

Associated with Q is an $n + 1$ -dimensional differential operator L :

$$L \equiv Q(d/dt, d/dy) = \sum c_{ij} (d/dt)^i (d/dy)^j.$$

Since L is determined by Q , the analytic properties of L evidently must follow from the algebraic properties of Q . The classification problem is the problem of finding correspondences between the analytic structure of L and the algebraic structure of Q .

In the classical cases, Q is quadratic, and a linear transformation of independent variables reduces L to a sum of squares, or to a single linear term plus a sum of squares. We obtain a classification for differential operators which corresponds to the classification of quadratic surfaces. If the number of independent variables is 2, we find exactly three “types” of second-order differential operators, which correspond to the three types of non-degenerate conic sections.

By a miraculous providence, to each of these three types of second-order equations there corresponds a famous problem of classical physics. Let $Q_H(a, b)$ denote the parabolic polynomial $a - b^2$. Then $L_H u \equiv u_t - y_{yy} = 0$ is the parabolic “heat equation,” which is satisfied by the temperature in a homogeneous heat conductor. The hyperbolic form $Q_W = a^2 - b^2$ is associated with the “wave equation,” $L_W u \equiv u_{tt} - u_{yy} = 0$, which describes small vibrations of an elastic medium. And if we interpret t as a spatial coordinate and let $Q_P = a^2 + b^2$, then the equilibrium state of a homogeneous medium (e.g., thermal equilibrium of a heat conductor) satisfies the “potential equation,” $L_P u \equiv u_{tt} + u_{yy} = 0$.

Reuben Hersh received his Ph.D. at NYU under Peter Lax. He has held positions at Fairleigh Dickinson, Stanford University, and the University of New Mexico. He spent a year leave at the Courant Institute. His main research is in Partial Differential Equations and in Random Processes. He has co-authored several recent articles in the *Scientific American*. *Editor.*

These physical interpretations help give us an orientation on the operator L . For instance, how many initial conditions should be given at $t = 0$ if $Lu = 0$ for positive t ?

Experience with ordinary differential equations suggests a number of independent initial conditions equal to l (the order of L with respect to d/dt). This answer is correct for the wave and heat equations, but not for the potential equation. This discrepancy is easy to understand in terms of the associated physical models. But from a mathematical viewpoint, the difference has to be explained by an algebraic difference between $a^2 + b^2$ on the one hand, and $a^2 - b^2$ or $a - b^2$ on the other.

DEFINITION 1. We will say **Cauchy's problem is correct for L** if, for arbitrary $f_k(y)$ which are in C_0^∞ (infinitely differentiable with compact support) there is a unique solution of

$$(2a) \quad Lu = 0 \text{ in } t > 0,$$

$$(2b) \quad (d/dt)^k u(0, y) = f_k, \quad 0 \leq k \leq l - 1.$$

We do not demand that the solution u be in C_0^∞ but only that it have as many derivatives as are required to be in the domain of Lu , and that these derivatives be in L_2 with respect to y , for each fixed t .

Correctness in this sense is an *intrinsic* property—a property of the solutions u of $Lu = 0$. We are asking for the corresponding *formal* property—an algebraic property of Q which should, in particular, be possessed by Q_H and Q_W but not by Q_P .

The three classical equations differ, not only in the number of conditions to give a well-posed problem, but also in the qualitative behavior of their solutions. The parabolic heat equation has a smoothing property: the solution u_H is infinitely differentiable, for positive t , even if the initial values f are discontinuous.

We can see this directly from the solution formula

$$u_H = (4\pi t)^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(\xi) \exp(-(y - \xi)^2/4t) d\xi.$$

DEFINITION 2. We say L is **intrinsically parabolic** if it is correct for Cauchy's problem even for initial data which are arbitrary members of $L_2(R^n)$, and if, moreover, the solution is infinitely differentiable.

What algebraic property of $Q_H = a - b^2$ corresponds to this intrinsic property of L_H ?

Solutions of the wave equation have an intrinsic property of great physical interest, the property of finite signal speed. By this we mean that if the initial data, f_0 and f_1 , have support in $|y| \leq M$ then, for some "signal speed" c , the solution at time t has support in $|y| \leq M + ct$.

We can see this directly from D'Alembert's formula,

$$u_W = \frac{1}{2} \left[f_0(y+t) + f_0(y-t) + \int_{y-t}^{y+t} f_1(s) ds \right].$$

In this case, $c = 1$.

DEFINITION 3. We say L is **intrinsically hyperbolic** if it is correct for Cauchy's problem and if, for some c , it has finite signal speed.

We have to find the formal property, possessed in particular by $Q_W = a^2 - b^2$, which corresponds to this intrinsic property.

Finally, for the potential equation, there is the property that the solution is determined, both for $t > 0$ and for $t < 0$, if its value is given for $t = 0$ (Dirichlet's problem). We shall use this to give a definition of **intrinsic ellipticity**, and seek a corresponding formal property.

We shall see that the Fourier transform answers all these questions with surprising ease and precision. Recall that if $f(y)$ is in $L_2(R^n)$ and

$$(3a) \quad \mathcal{F}f \equiv \hat{f}(\eta) = (2\pi)^{-n/2} \int \exp(-i\eta y) f(y) dy_1 \cdots dy_n,$$

then $\hat{f}(\eta)$ is in $L_2(R^n)$ and

$$(3b) \quad \mathcal{F}^{-1}\hat{f} \equiv f(y) = (2\pi)^{-n/2} \int \exp(i\eta y) \hat{f}(\eta) d\eta_1 \cdots d\eta_n.$$

For $u(t, y)$, we denote by $\mathcal{F}u \equiv \hat{u}(t, \eta)$ the transform in y alone, t remaining constant. If f and all its derivatives are in L_2 , integration by parts gives

$$(4) \quad \mathcal{F}((d/dy)^k f) = (i\eta)^k \hat{f}(\eta).$$

Since \mathcal{F} is a linear operator, we get, combining (3) and (1),

$$(5) \quad \begin{aligned} \mathcal{F}[Lu] &= \mathcal{F}[Q(d/dt, d/dy)u] = \mathcal{F} \sum c_{ij}(d/dt)^i (d/dy)^j u \\ &= \sum c_{ij}(d/dt)^i (i\eta)^j \hat{u} = Q(d/dt, i\eta) \hat{u}. \end{aligned}$$

From (4) we see also that if df/dy exists and is in L_2 , then $\eta \hat{f}$ is in L_2 . In other words, the smoother f , the more rapid the decay of \hat{f} at infinity. In the same way, using (3b), we see that the more rapidly f decays, the smoother is \hat{f} .

2. Equations correct for Cauchy's problem. Now we are ready to use the Fourier transform to classify Q . The first problem is to identify Q for which Cauchy's problem (2a, b) is correct. In view of (5) \hat{u} satisfies

$$(6a) \quad Q(d/dt, i\eta) \hat{u} = 0,$$

$$(6b) \quad (d/dt)^k \hat{u}(0) = \hat{f}_k, \quad 0 \leq k \leq l-1.$$

Equation (6a) is an ordinary differential equation in t . The coefficients are functions of the parameter η , but do not depend on t , so for each η it is solved in the same

explicit elementary way as an ordinary differential equation with constant coefficients. One simply writes down the general solution

$$(7) \quad \hat{u} = \sum_r \sum_s^{m_r} \gamma_{rs}(\eta) t^s \exp(\tau_r(\eta)t),$$

where $\tau_r(\eta)$ is a root of $Q(\tau, i\eta) = 0$ with multiplicity $m_r - 1$. The coefficients γ_{rs} are found by solving the linear algebraic system of equations which results from substituting (7) into (6b).

In the case of the wave equation $Q_W = a^2 - b^2$, we have $l = 2$ and we find

$$(8) \quad \begin{aligned} \hat{u}_W &= \hat{f}_0 \cos \eta t + \hat{f}_1 \frac{\sin \eta t}{\eta} \\ &= \frac{1}{2} \hat{f}_0 (e^{i\eta t} + e^{-i\eta t}) + \frac{1}{2i\eta} \hat{f}_1 (e^{i\eta t} - e^{-i\eta t}). \end{aligned}$$

For the heat equation, with $l = 1$, we get

$$(9) \quad \hat{u}_H = e^{-\eta^2 t} \hat{f}_0.$$

In both cases, it is clear that the solution-transforms \hat{u} decay, as $|\eta| \rightarrow \infty$, at least as rapidly as the data-transforms \hat{f} , since the exponential multipliers $e^{\pm i\eta t}$ and $e^{-\eta^2 t}$ are bounded by 1. Now, assuming that the data f are infinitely differentiable, with all derivatives in L_2 , it follows from (4) that the data-transforms \hat{f} decay more rapidly than $|\eta|^{-r}$ for any r . Then \hat{u}_H and \hat{u}_W also decay more rapidly than $|\eta|^{-r}$. From this it follows that the inverse transform (3b) converges uniformly and absolutely, and the solutions u_H and u_W exist and are infinitely differentiable, with all derivatives in L_2 .

On the other hand, what happens if we try to solve (2) for the potential equation, $Q_P = a^2 + b^2$?

This time we get for \hat{u} the formula

$$(10) \quad \hat{u}_P = \frac{1}{2} e^{\eta t} \left(\hat{f}_0 + \frac{\hat{f}_1}{\eta} \right) + \frac{1}{2} e^{-\eta t} \left(\hat{f}_0 - \frac{\hat{f}_1}{\eta} \right).$$

If the solution u exists, it is to be got from (10) by means of the inversion integral (3b). But the factor $e^{\eta t}$ in the first term of (10) will certainly make the inverse transform integral diverge for $t > 0$, unless $\hat{f}_0 + \hat{f}_1/\eta$ decays as rapidly as $e^{-\eta t}$. This is a severe constraint on the choice of f_1 , given f_0 . On the other hand, if we prescribe only $u(0, y) = f_0$ (Dirichlet's problem) then (6b) consists of only a single equation. It is satisfied by

$$(11) \quad \hat{u}_P = e^{-t|\eta|} \hat{f}_0$$

which, for $t \geq 0$, clearly has an inverse transform u . (The problem for $t < 0$ would be solved by $e^{t|\eta|} \hat{f}_0$.)

Thus we can see directly from examination of the transforms, without actually finding any of the solutions u_H, u_W or u_P , that Cauchy's problem is indeed correct for $a - b^2$ and $a^2 - b^2$, but not for $a^2 + b^2$. These examples suggest that the number of conditions we can impose at $t = 0$ is equal to the number of exponential terms in (7) which are bounded as $|\eta| \rightarrow \infty$. Altogether there are l terms, as many as the order in d/dt of L . In Cauchy's problem we have to satisfy l conditions, so we need to have all of them well-behaved. We are led to

DEFINITION 4. Q is **correct in the sense of Petrowsky** if, for all real n -tuples η , all roots $\tau_f(\eta)$ of $Q(\tau, i\eta) = 0$ satisfy

$$(12) \quad \operatorname{Re} \tau \leq M \text{ for some constant } M.$$

We claim that if Q is correct in the sense of Petrowsky, then Cauchy's problem is correct for L . We need only show that \hat{u} as given by (7) exists and decays more rapidly than $|\eta|^{-r}$ for any r . (The data f are assumed to be infinitely differentiable with compact support.)

As to the existence of \hat{u} , all we need to do is find the $\gamma_{r,s}(\eta)$ by solving the linear algebraic equations,

$$\left(\frac{d}{dt}\right)^k \left(\sum_{r,s} \gamma_{rs} t^s e^{t\tau_r}\right)\Big|_{t=0} = 1 \hat{f}_k \quad \text{for } k = 0, \dots, m-1.$$

If the roots τ are distinct, so that no powers of t actually appear in front of the exponential, then the coefficient matrix of this system is the familiar Vandermonde matrix, which is known to be nonsingular so long as $\tau_j \neq \tau_k$ for $j \neq k$. In case of multiple roots, we get a closely related "generalized Vandermonde matrix." The reader who has never seen it done will find it an enjoyable exercise in old-fashioned advanced algebra to show that in this case also the matrix is non-singular. This takes care of the first point, existence of \hat{u} for all finite η .

To show that \hat{u} decays rapidly at $|\eta| \rightarrow \infty$, we have to consider the dependence on η of these coefficients $\gamma_{rs}(\eta)$ whose existence we have just established. Since they are solutions of a system of linear equations whose coefficients are powers of η and whose right sides are the data-transforms $\hat{f}_j(\eta)$, Cramer's rule for solving linear equations shows that $\gamma_{jk}(\eta)$ is less than some polynomial in (η) times $\sup_j |\hat{f}_j(\eta)|$.

If f_j is in C_0^∞ , $\hat{f}_j(\eta)$ is $O(|\eta|^{-n})$ for any n . The same is then true of $\gamma_{rs}(\eta)$, and therefore, by virtue of (7) and (12), of \hat{u} as well, as we claimed.

It is natural to ask whether the converse is also true, whether it is necessary for Q to be Petrowsky-correct in order for L to be correct for Cauchy's problem. The answer is yes. This is not obvious on the face of it. From (7) it would seem that all one needs for existence of u is

$$(13) \quad \operatorname{Re} \tau_r(\eta) \leq M \log |\eta|.$$

But it can be shown that (13) and (12) are equivalent. This follows from an important algebraic lemma:

LEMMA. *The function $\Lambda(z)$ defined by*

$$\Lambda = \sup_{|\eta|=z} \operatorname{Re} \tau_r(\eta), \quad Q(\tau, i\eta) = 0$$

is a piecewise algebraic function of z .

This lemma is a special case of a famous theorem of Tarski on the decidability of the elementary theory of real polynomials.

Tarski's theorem says that one can eliminate (solve for) some of the variables from a given system of real polynomial equations and inequalities, provided the remaining variables satisfy one of a finite number of finite systems of polynomial equations and inequalities. A short and elegant proof of Tarski's theorem has been given by Paul Cohen [10]. A full discussion, with an exposition of Seidenberg's proof of Tarski's theorem, is in [8].

Here we simply point out the implications of the lemma for our classification problem. Near $|z| = \infty$, $\Lambda(z) = \Lambda(|\eta|)$ is equal to one of a finite collection of algebraic functions of $|\eta|$. Each of these algebraic functions of one variable has a Puiseux expansion in fractional powers of $|\eta|$. (No such expansion exists for algebraic functions of two or more variables. It is precisely to make use of this one-variable result that the lemma is necessary.) We conclude that if any of the roots violates (12), then it must go to $+\infty$ like some positive fractional power of $|\eta|$, and so it must also violate (13). Thus we see that Petrowsky-correctness of Q is indeed equivalent to intrinsic correctness of L .

3. Parabolic equations. Next we define "parabolic." We mentioned earlier that u_H , the solution of Cauchy's problem for the heat equation, is infinitely differentiable even for singular initial data. We now can read off this conclusion from (11). Indeed, if f_0 is in L_2 , then so is \hat{f}_0 . The factor $e^{-\eta^2 t}$ makes $\hat{u}(t, \eta)$ decay more rapidly than any negative power of $|\eta|$, which means that $u(t, y)$ is infinitely differentiable.

This pleasant state of affairs comes about because the root τ_H of the equation $\tau - (i\eta)^2 = 0$ goes to $-\infty$ as $|\eta| \rightarrow \infty$. We are thereby led to

DEFINITION 5. Q is **formally parabolic** if all roots τ_j of $Q(\tau, i\eta) = 0$ satisfy

$$(14) \quad \operatorname{Re} \tau \rightarrow -\infty \text{ as } |\eta| \rightarrow \infty \text{ through real values of } \eta.$$

We want to prove that if Q is formally parabolic, then L is intrinsically parabolic. Since formal parabolicity implies Petrowsky-correctness, we need only prove that $u(t, y)$ is infinitely differentiable even for arbitrary L_2 initial data. We again refer to (7). We have seen that $\gamma_{rs}(\eta)$ grow at most like some power of $|\eta|$. We need to show that \hat{u} decays faster than $|\eta|^{-n}$ for arbitrary n . Now, by the same algebraic argument we used above, one can show that if $\operatorname{Re} \tau_r$ goes to $-\infty$ as $|\eta| \rightarrow \infty$, then for some

positive constants α and β , $\operatorname{Re} \tau_r \leq -\alpha |\eta|^\beta$. This estimate does guarantee that \hat{u} decays faster than $|\eta|^{-r}$ for any r , and so u is infinitely differentiable, as we wanted to prove.

To prove the converse, one shows that if some root τ of $Q(\tau, i\eta) = 0$ fails to satisfy (14), then, by an appropriate choice of data f_k one can construct a solution $u(t, y)$ which is not C^∞ , that is, whose transform \hat{u} does not decay like $|\eta|^{-n}$ for some n .

It should be emphasized that in this definition, unlike a usage sometimes found in the older literature, an equation by no means need be parabolic merely because it is of lower order in some variables than in others. For example, the equation $u_t = iu_{xx}$ is *not* parabolic, as the reader may verify.

From the intrinsic form of the definition, it is immediate, without any computation, that any smooth invertible transformation of the y -variables takes a parabolic equation into a parabolic equation. The same is true for any smooth transformation of the closed half-line $0 \leq t < \infty$ onto itself.

If Q is parabolic, then it is actually the case that any solution of $Lu = 0$ is infinitely differentiable, even if u is defined only on some open subset of $t-y$ space. This property is called "hypoellipticity." A famous result of Hormander is that L is hypoelliptic if and only if, for all complex η , $Q(\eta) = 0$ and $|\eta| \rightarrow \infty$ imply $|\operatorname{Re} \eta| \rightarrow \infty$. (Here, unlike Definition 5, there is no "distinguished" variable τ . All the variables are treated as equals.)

4. Hyperbolic equations. Next, we look for an algebraic criterion for hyperbolicity. We must find a property of \hat{u} which is related to finite signal speed of u in the same way that rapid decay of \hat{u} is related to smoothness of u . This need is precisely filled by the theorem of Paley and Wiener.

Suppose the data f in (7) vanish for $|y| > M$. Then the domain of integration in (3) is a bounded set, and it is clear that the integral converges even if η is complex. Moreover, the resulting function $\hat{f}(\eta)$ is differentiable, for all real or complex η , and it satisfies the obvious estimate,

$$(15) \quad |\hat{f}(\eta)| \leq \text{const } e^{M|\eta|}.$$

The constant is just the maximum value of $|f(y)|$. In brief, \hat{f} is entire analytic, of exponential type M . For real η , \hat{f} is of course in L_2 .

The Paley-Wiener theorem says that, conversely, if \hat{f} is entire, satisfies (15), and is square-integrable for real η , then $f(y) \equiv 0$ for $|y| > M$.

A more precise version has been given by Plancherel and Polya. Given a closed bounded set \mathcal{D} in real y -space, the "support function" of \mathcal{D} is defined by $s(y) = \max_{x \text{ in } \mathcal{D}} x \cdot y$. If \hat{f} is square-integrable for real η , and if \hat{f} is entire analytic and satisfies, for all complex η ,

$$|\hat{f}(\eta)| \leq \text{const } \exp |s(\operatorname{Im} \eta)|$$

then f vanishes outside \mathcal{D} .

The idea of the proof is simple. One considers an arbitrary y not in \mathcal{D} , and chooses w so that $y \cdot w > s(w)$. In the inversion formula (3b), since the integrand is now entire analytic, we may shift the path of integration into complex η -space, taking $\eta + kw$, $-\infty < \eta < \infty$, as the new path of integration. By Cauchy's theorem, the value of the integral is unchanged. If k is very large, $|\hat{f}|$ will be as small as we please.

Let us apply this theorem to the wave equation. First we make the obvious remark that, if two functions are entire of exponential type M_1 and M_2 , then their product is also entire, of exponential type $M_1 + M_2$. Now, if the data f_0 and f_1 have support in $|y| < M$, then, by the converse of the Paley-Wiener theorem, their transforms \hat{f}_0 and \hat{f}_1 are entire functions of exponential type M . Since $\cos \eta t$ and $\sin \eta t / \eta$ are, as functions of η , entire of exponential type t , we see from formula (8) that \hat{u}_w is entire of exponential type $M + t$, and so by Paley-Wiener, u_w , the solution of Cauchy's problem for the wave equation, vanishes for $|y| > M + t$. That is, the wave equation has finite speed of propagation.

For the general case, we have to look at (7). Now we suppose that $f_k(y) \equiv 0$ for $|y| > M$, so all the data-transforms \hat{f}_k are entire of exponential type M . It follows, without any additional hypotheses, that \hat{u} is an entire analytic function of η . To see this, we recall that by our previous analysis, the functions $\gamma_{rs}(\eta)$ are equal to linear combinations of \hat{f}_k , with coefficients which are rational functions of η and of the roots $\tau_r(\eta)$. Now, the roots τ_r are multivalued algebraic functions of η , so there seems to be a question about the analyticity of \hat{u} at branch points of τ . But even though each root τ_r is multivalued, \hat{u} is continuous and singlevalued as a function of real or complex η ; indeed, \hat{u} is the solution of the ordinary differential equation (6a, b), and so is single-valued and depends continuously on the parameter η . Since \hat{u} is continuous and single-valued for all complex η , and in particular, at the branch points of $\tau(\eta)$, it is entire analytic in each η_j separately, and therefore it is an entire analytic function of the n -tuple (η_1, \dots, η_n) .

What is needed is an extra hypothesis on Q to make \hat{u} a function of exponential type.

DEFINITION 6. We say Q is **formally hyperbolic** if it is correct in the sense of Petrowsky and if, in addition, the roots $\tau(\eta)$ of $Q(\tau, i\eta) = 0$ satisfy, for all complex η , $|\operatorname{Re} \tau(\eta)| \leq C_0 |\eta| + C_1$, for some constants C_0 and C_1 .

If Q satisfies Definition 6, and if the data-transforms are entire of exponential type M , then \hat{u} is entire of exponential type $tC_0 + M$. The same reasoning as in the special case of the wave equation now shows that if Q is formally hyperbolic, L is intrinsically hyperbolic, with signal speed not greater than C_0 . The converse is also true, and is proved with the help of our algebraic lemma.

An important example of a hyperbolic equation is the first-order equation,

$$(16) \quad Qu \equiv u_t - \sum a_j \frac{\partial u}{\partial y_j} = 0.$$

If we let u be a vector and a_j be symmetric matrices, (16) becomes a symmetric hyperbolic system of equations. It is hyperbolic because the roots $\tau(\eta)$ of $\det Q(\tau, i\eta) = 0$ are simply the eigenvalues of the skew-Hermitian matrix $i \sum \eta_j a_j$, so they have zero real part and grow like a first power of $|\eta_j|$ for each j . Maxwell's equations in vacuo are a famous example of a symmetric hyperbolic system.

If the degree of $Q(a, b)$ as a polynomial in a and b is the same as l , its degree in a alone, one says that Q is "non-characteristic" for the initial plane $t = 0$. By comparing Definitions (4) and (6), it is easy to see that Q is formally hyperbolic if and only if it is noncharacteristic and Petrowsky correct. In view of the equivalence we have shown between the formal and intrinsic properties, this means that for non-characteristic polynomials, correctness for Cauchy's problem (in the sense of Definition 1) implies finite signal speed. (By contrast, the heat equation and other parabolic equations are characteristic for t .)

On the other hand, the finite signal speed property actually implies that Cauchy's problem is correct, even for initial data which grow at $|y| \rightarrow \infty$ with arbitrary rapidity. To prove this, one writes the initial data as a sum of terms with support in the cells, $j \leq y_j \leq j + 1$. To each term in this expansion there corresponds a term in the solution; each term has compact support and can be obtained by the Fourier transform method.

By the finite signal speed property, all but a finite number of these solution terms must vanish for any given value of t , in any cell $j \leq y_j \leq j + 1$. Therefore, the sum exists and is the solution to the given Cauchy problem with arbitrary unbounded data.

An interesting difference between hyperbolicity and parabolicity appears if we consider problems in a region with boundary. Suppose that $Lu = 0$ is satisfied only for $t > 0$, $y_1 > 0$, and on $y_1 = 0$, $t > 0$, u satisfies some set of boundary conditions, $B_j(d/dt, d/dy)u = 0$. If L is parabolic, then so long as a solution exists, it will be smooth in the interior, ($t > 0$, $y_1 > 0$); the choice of boundary operators cannot overcome the smoothing property of L . But if L is hyperbolic, the finite signal speed of L may be lost if surface waves are propagated with infinite speed along the boundary $y_1 = 0$.

The simplest example is to take

$$Lu \equiv u_{tt} - u_{y_1 y_1} - u_{y_2 y_2} = 0 \text{ in } t > 0, \quad y_1 > 0$$

with the single boundary condition

$$(17) \quad Bu \equiv u_t - u_{y_2 y_2} = 0 \text{ on } y_1 = 0.$$

We also need, of course, initial values, $u(0, y_1, y_2) = f_0(y_1, y_2)$, $u_t(0, y_1, y_2) = f_1(y_1, y_2)$ subject to the compatibility condition $f_1(0, y_2) = \partial^2 / \partial y_2^2 f_0(0, y_2)$.

Since the boundary condition involves only tangential derivatives, one can find the solution u on the boundary $t > 0$, $y_1 = 0$, by solving Cauchy's problem for the

heat equation (17) with initial value $f_0(0, y_2)$. Then u is determined in the interior, $t > 0$, $y_1 > 0$, by solving a standard boundary value problem, where u itself is now known on the boundary $y_1 = 0$. Now even if the initial data vanish except near the origin, $y_1 = y_2 = 0$, it is clear that because the boundary operator B is not hyperbolic, the solution u will in general be non-zero for arbitrarily large y_2 if $y_1 \leq t$. Thus the finite speed of propagation is lost.

To guarantee finite speed of propagation for a mixed initial-boundary value problem, one needs to impose a hyperbolicity condition on the boundary operators B_j as well as the interior operator $L = Q(D)$.

To find such a condition, one again uses an integral transform to reduce the partial differential equation to an ordinary differential equation. Fourier transform can again be applied in all the y_j except y_1 , the variable which has a boundary. If we then use a Laplace transformation in t , we again end up with an ordinary differential equation, but now in y_1 instead of in t .

Again the Paley-Wiener theorem yields the precise algebraic condition corresponding to finite signal speed (see [7]). A simple sufficient condition, in addition to unique solvability of the mixed initial boundary value problem, is that L is hyperbolic and B_j is homogeneous as a polynomial in d/dt and d/dy .

It should be mentioned that despite the false impression created by the traditional classification, Cauchy's problem is correct for many operators which are neither parabolic nor hyperbolic. A few which have physical interest appear in the Schrödinger equation $u_t = iu_{yy}$, the vibrating beam equation $u_{tt} + u_{yyyy} = 0$, and the equation of viscous acoustics,

$$u_{tt} = 2u_{t_{yy}} + u_{yy}.$$

5. Elliptic equations. We now have to consider those operators for which Cauchy's problem is *not* well posed. This means there are some "badly behaved" roots $\tau(\eta)$ whose real part goes to $+\infty$ as $|\eta| \rightarrow \infty$ through imaginary values. It is clear that reasoning identical to that we have used before would show that we can prescribe a number of initial conditions for u just equal to the number of "well behaved" roots τ_k of $Q(\tau, i\eta) = 0$ that satisfy (12).

This sheds some light on what is wrong with an "ultra-hyperbolic" equation such as

$$u_{tt} = u_{y_1 y_1} - u_{y_2 y_2}.$$

For this example, the roots $\tau(\eta)$ have the form $\tau_{\pm} = \pm(\eta_2^2 - \eta_1^2)^{\frac{1}{2}}$. As η_1 and η_2 go to infinity, $\text{Re } \tau_{\pm}$ may remain bounded or may go to $\pm\infty$, depending on our path in the η_1 - η_2 plane.

Thus, we are unable to prescribe in a natural way a definite number of conditions for u at $t = 0$. The number of conditions k should equal the number of roots $\tau_j(\eta)$

satisfying (12), but this number is not well-defined. Shilov has met this difficulty by prescribing the initial conditions in terms of the transform \hat{u} ; then indeed one can formulate appropriate initial conditions for every Q . Those Q for which k is independent of η are called “regular” by Shilov.

For arbitrary Q , not necessarily regular, and for each path θ which goes to ∞ in real η -space, we can define $k(\theta)$ as the number of roots of

$$Q(\tau, i\eta) = 0 \text{ which satisfy } \operatorname{Re} \tau(\eta) < C \text{ on } \theta.$$

Then there will always be two integers,

$$k_1 = \min_{\theta} k(\theta) \text{ and } k_2 = \max_{\theta} k(\theta).$$

It takes only a moment's reflection to see that k_1 is the number of conditions we can prescribe arbitrarily in L_2 at $t = 0$. On the other hand, if we have a solution u which satisfies k_2 prescribed conditions at $t = 0$, then its transform \hat{u} (and so u itself) are thereby uniquely determined. Thus the “regular” case—the case $k_1 = k_2$ —is precisely the case where it is possible to obtain *both* existence and uniqueness by prescribing initial conditions in a manner independent of η —i.e., in terms of the “physical” variables t and y .

By applying these remarks on “regular” boundary problems, we can generalize the half-space Dirichlet's problem for Laplace's equation. We will consider only real polynomials Q which are homogeneous of even order $2k$.

We now use $x = 0$ instead of $t = 0$ to specify our boundary. We suppose that $Q(d/dx, d/dy)$ is a real, homogeneous polynomial of even order $2k$; $y = (y_1, \dots, y_n)$ is an n -vector for some $n \geq 1$.

DEFINITION 7. We shall call Q **intrinsically elliptic** if we can solve $Q(D)u \equiv Lu = 0$ both in $x > 0$ and in $x < 0$, with k boundary conditions,

$$\left. \frac{\partial^r u}{\partial x^r} \right|_{x=0} = f_r \quad \text{for } r = 0, \dots, k-1,$$

for “arbitrary” f_r in L_2 . (The reader should be warned that this definition is not standard, though it is consistent with the standard one. We shall discuss this point below.)

To get a corresponding “formal” definition, we notice the most obvious property of the Laplace polynomial $Q(a, b) = a^2 + b^2$ —it is positive in the real (a, b) -plane except at the origin.

DEFINITION 8. A homogeneous, real polynomial Q is **formally elliptic** if $Q(a, b) \neq 0$ for real a, b not both zero.

Because of the homogeneity of Q , the roots $\xi(\eta)$ of $Q(\xi, i\eta) = 0$ satisfy $\xi(s\eta) = s\xi(\eta)$ for all complex s . Furthermore, for real η , $\xi(\eta)$ is not pure imaginary, for then we would have $Q(\xi, i\eta) = i^{2k}Q(\xi/i, \eta) = 0$ with real ξ/i and η , violating the formal

ellipticity. Because the coefficients are real, the roots ξ/i of $Q(\xi/i, \eta) = 0$, which are all non-real, for real η , consist of pairs of conjugate complex numbers. Counting them according to multiplicity, there must be k each in the upper and lower half plane, so that there are exactly k of the roots $\xi(\eta)$ which satisfy (12), and k which satisfy the opposite inequality. These remarks are enough to show that there exists a unique L_2 solution to our generalized "Dirichlet problem" for either $x > 0$ or $x < 0$. That is, our version of "intrinsic ellipticity" follows from formal ellipticity. Now, the homogeneity implies

$$\xi(\eta) = |\eta| \xi\left(\frac{\eta}{|\eta|}\right) \text{ so that } |\operatorname{Re} \xi| \rightarrow \infty, \text{ as } |\eta| \rightarrow \infty \text{ for each root } \xi(\eta).$$

Suppose $\eta = (\eta_1, \dots, \eta_n)$ and $y = (y_1, \dots, y_n)$ are n -tuples and ηy means $\sum \eta_i y_i$. Then the n -fold iterated integral

$$u(x, y) = \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} e^{-i\eta y} \sum_{j,k} \gamma_{jk}(\eta) x^j e^{x \xi_k(\eta)} d\eta$$

can be written as

$$\sum_{j,k} \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} x^j \gamma_{jk}(\eta) \exp\left(-i\eta y + |\eta| x \xi_k\left(\frac{\eta}{|\eta|}\right)\right) dy.$$

Suppose, for definiteness, we are interested in the case $x > 0$. Then we choose ξ_k so that $\operatorname{Re} \xi_k < 0$. Since $\operatorname{Re} \xi_k(\eta/|\eta|)$ is then a continuous negative function on a compact set (the unit sphere in real η -space), it assumes a negative maximum. We can therefore choose a number Λ independent of k , such that $\operatorname{Re} \xi_k(\eta/|\eta|) \leq \Lambda < 0$, for all real η and for all the ξ_k with negative real parts. Then it is easy to see that the integral for $u(x, y)$ converges uniformly for complex x and y such that $|\operatorname{Im} y| < \operatorname{Re} x / \Lambda$, so that, as in the classical second-order case, $u(x, y)$ is real-analytic in both x and y .

I. G. Petrowsky proved (1938) that if $Q(d/dx, d/dy)$ is formally elliptic, then all solutions of $Lu = 0$ are analytic (not just those defined in a half-space $x > 0$, which we are considering). This property is the one usually taken as the definition of intrinsic ellipticity. We have deviated from this tradition for the sake of simplicity and brevity.

We are now in a position to tie together the notions of parabolic and elliptic. Consider $Q(d/dt, d/dy) = d/dt - Q_0(d/dy)$, where Q_0 is homogeneous of order $2k$, real and elliptic, and y is an n -tuple, $n \geq 2$. Then the roots $\tau(\eta)$ of $Q(\tau, \eta) = 0$ are just

$$\begin{aligned} \tau(\eta) &= Q_0(i\eta) = i^{2k} Q_0(\eta) \\ &= (-1)^k |\eta|^{2k} Q_0\left(\frac{\eta}{|\eta|}\right). \end{aligned}$$

Now $Q_0 \neq 0$ on the unit η -sphere, and since η , like y , is an n -tuple, $n \geq 2$, the unit sphere is a compact connected set. It follows that either

$$Q_0 \left(\frac{\eta}{|\eta|} \right) \leq -C < 0 \text{ or } Q_0 \left(\frac{\eta}{|\eta|} \right) \geq C > 0,$$

for all real η . We can strengthen our definition of formal ellipticity to require

$$\operatorname{sgn} Q_0(\eta) = (-1)^{k+1} \text{ for real } \eta.$$

Then it is clear that $\operatorname{Re} \tau(\eta) = Q_0(i\eta)$ goes to $-\infty$ as $|\eta| \rightarrow \infty$ through real values, and $d/dt - Q_0(d/dy)$ is parabolic.

This discussion of problems with constant coefficients in a half-space is, of course, only a bare introduction to some of the concepts and problems of linear differential operators. It is natural, for example, to ask what happens if our domain is more complicated, or if our equation has variable coefficients. Roughly speaking, one might hope that if the coefficients do not vary too wildly, then an equation with variable coefficients is intrinsically parabolic (or hyperbolic or elliptic) if it is formally of that type at each point. This expectation turns out to be pretty generally true. The elliptic case is the best known, and the one for which the least regularity of the coefficients need be assumed. (Strong existence and regularity theorems have been obtained where the coefficients are no more than measurable as functions of position!) Usually some sort of uniformity must be imposed in the ellipticity (or parabolicity or hyperbolicity) of the variable coefficients. Then it is often possible to treat variable coefficient problems as perturbations of constant-coefficient problems.

A different direction in which one can generalize the classification problem is to replace the differential operator d/dy by an abstract operator A . If A generates a group, it turns out, rather unexpectedly, that the classification of polynomials $Q(d/dt, A)$ can be reduced to the classical case $Q(d/dt, d/dy)$. (See [9].)

Two standard references on linear differential operators are Hormander [1] and Gel'fand and Shilov [2], especially volume 3. The book by Shilov [3] has an extensive discussion on half-space problems. The books by Treves [4] and Friedman [5] should also be consulted. For hyperbolic equations Lax's notes [6] are outstanding.

References

1. L. Hormander, *Linear Partial Differential Operators*, Academic Press, New York, 1963.
2. I. M. Gel'fand and G. E. Shilov, *Generalized Functions*, (Vol. 3), Academic Press, New York, 1967.
3. G.E. Shilov, *Generalized Functions and Partial Differential Equations*, Gordon and Breach, New York 1966.
4. F. Treves, *Linear Partial Differential Equations with Constant Coefficients*, Gordon and Breach, New York, 1966.
5. A. Friedman, *Generalized Functions and Partial Differential Equations*, Prentice-Hall, 1963.
6. Peter Lax, *The Theory of Hyperbolic Equations*, Lectures Notes, Stanford University.

7. R. Hersh, Boundary conditions for equations of evolution, Arch. Rational Mech. Anal. Vol. 16, No. 4. (1964) 243-264.

8. E. A. Gorin, Asymptotic properties of polynomials and algebraic functions of several variables, Uspekhi Mat. Nauk (N.S.) 16, No. 1, 93-119. Russian Mathematical Surveys, 1961.

9. R. Hersh, Explicit solution of a class of higher-order abstract Cauchy problems, J. Differential Equations, 8 (1970) 570-579.

10. Paul J. Cohen, Decision procedures for real and p -adic fields, Comm. Pure Appl. Math., 22 (1969) 131-151.

THE CESÀRO OPERATORS AND THEIR GENERALIZATIONS: EXAMPLES IN INFINITE-DIMENSIONAL LINEAR ANALYSIS

GERALD LEIBOWITZ, University of Connecticut

In this article I shall describe some recent results concerning a class of linear transformations which were originally studied many years ago by the great Felix Hausdorff from the point of view of summability theory (that is, as generalized averaging processes) and which have come to be called Hausdorff transformations. I shall also indicate some of the methods used in obtaining these results.

The Hausdorff transformations, of which the familiar methods of Cesàro means are the simplest examples, provide a good illustration of the extent to which linear operators on infinite-dimensional spaces share some of the properties of their finite-dimensional counterparts and of the extent to which their behavior can be very different. We shall see an interesting interplay among concepts and techniques from basic linear algebra and calculus, classical analysis, and abstract analysis.

1. Preliminaries. Recall that a **Banach space** is a normed linear space X in which every Cauchy sequence is convergent, or equivalently, in which every absolutely summable series is summable ($\sum \|x_n\| < \infty$ implies that $\sum x_n$ exists in X). We shall be concerned primarily with the Banach spaces defined as follows. Let p be a real number greater than 1. The space l^p consists of all sequences $s = \{s_n: n = 0, 1, 2, \dots\}$ such that $\|s\|_p = (\sum |s_n|^p)^{1/p}$ is finite. The space $L^p(0, 1)$ consists of all Lebesgue measurable functions f on the unit interval for which $|f|^p$ is integrable, with $(\int_0^1 |f|^p)^{1/p}$ as norm, and the space $L^p(0, \infty)$ is the analogous space of functions on the positive real axis. (Throughout this article, all sequences and functions are taken as complex-valued, and "scalar" will mean "complex number".)

By an **operator** on a Banach space X we shall mean a continuous linear trans-

Gerald Leibowitz received his M. I. T. doctoral degree under Kenneth Hoffman. He held a position at Northwestern University before assuming his present post. He was Associate Director of CUPM in 1968/69 and is presently a member of the Consultants Bureau. His work on functional analysis includes *Lectures on Complex Function Algebras* (Scott, Foresman and Co., 1970). *Editor.*

formation from X into itself. The operators on X again form a Banach space if the relevant structures are defined as follows:

$$(S_1 + S_2)(x) = S_1x + S_2x, (\alpha S)(x) = \alpha \cdot Sx, \|S\| = \sup \{\|Sx\| : \|x\| \leq 1\}.$$

Associated with each operator S are two subsets of the complex plane. The **resolvent set** for S is the set $\rho(S)$ consisting of all scalars λ such that $\lambda I - S$ is invertible, where I is the identity transformation on X . (**Invertibility** here means that the operator is one-to-one and has full range; continuity of the inverse transformation then follows from Banach's closed graph theorem. See [13, p. 236].) The **spectrum** of S , denoted by $\sigma(S)$, is the complement of $\rho(S)$.

The spectrum of an operator on X is a generalization of the corresponding finite-dimensional notion. Indeed, if $\dim X < \infty$, then $\sigma(S)$ is precisely the set of eigenvalues of S . However, if X is of infinite dimension, an operator need not have any eigenvalues, and if it does, they need not exhaust its spectrum. (We shall see examples below.) Much is known about spectra of operators. First, if $|\lambda| > \|S\|$, then $\sum_0^\infty \lambda^{-1-n} S^n$ converges in the operator norm to the inverse of $\lambda I - S$, hence $\sigma(S)$ is contained in the disk with radius $\|S\|$ centered at the origin. Moreover, the inverse of an operator is a continuous function of the operator (see [13], p. 306), so $\rho(S)$ is an open set. Thus the spectrum of every operator is closed and bounded, hence compact. An application of Liouville's theorem to the operator-valued analytic function $(\lambda I - S)^{-1}$ reveals that $\rho(S)$ is a proper subset of the plane, so $\sigma(S)$ is not empty. (On the other hand, each nonvoid compact set of scalars is the spectrum of some operator.)

We end this section with a brief discussion of dual spaces and operator adjoints. Given a Banach space X , the **dual space** X^* consists of all continuous linear functions from X to the scalar field, furnished with the following structures:

$$(\phi + \psi)(x) = \phi(x) + \psi(x), (\alpha\phi)(x) = \alpha\phi(x), \|\phi\| = \sup \{|\phi(x)| : \|x\| \leq 1\}.$$

X^* is again a Banach space. (One can identify the duals of l^p , $L^p(0,1)$, and $L^p(0,\infty)$ with the spaces l^q , $L^q(0,1)$, $L^q(0,\infty)$ where $1/p + 1/q = 1$. The dualities are defined as follows: $\langle s, t \rangle = \sum s_n t_n$, s in l^p , t in l^q ; $\langle f, g \rangle = \int f(x)g(x)dx$, f in L^p , g in L^q .) Each operator S on X determines an operator S^* on X^* , the **adjoint** of S , given by $(S^*\phi)(x) = \phi(Sx)$. The mapping $S \rightarrow S^*$ is linear, preserves operator norms, and is anti-multiplicative: $(S_1 S_2)^* = S_2^* S_1^*$. From this it is clear that $\rho(S)$ is contained in $\rho(S^*)$, or equivalently that the spectrum of S^* is a subset of the spectrum of S . If, moreover, each continuous linear functional on X^* is given by evaluation at some member of X (in this case one writes $X = X^{**}$ and says that the space X is reflexive), then S and S^* have the same spectrum. This is an important remark for our study below since we shall determine the spectra of certain operators in two stages: in the first it is shown that one of S or S^* has a certain large set of eigenvalues; in the second it is shown that for structural reasons, the spectrum cannot contain any scalars outside the closure of the eigenvalue set.

2. Cesàro operators. Associated with each sequence $s = \{s_n\}$ of complex numbers is its sequence of arithmetic means $\{\pi_n\}$ given by

$$\pi_n = (s_0 + s_1 + \cdots + s_n)/(n+1), \quad n = 0, 1, 2, \dots$$

The basic theorem of Hölder and Cesàro asserts that if s_n converges to a limit s_∞ , then $\pi_n \rightarrow s_\infty$ as well. Any sequence-to-sequence transformation which preserves convergence and limits of all convergent sequences is said to be a **regular method of summability**. In the next section we shall consider the properties of a large class of summability methods — which contains many of the best-known regular methods — using the Hölder-Cesàro operator as our model. (For information about summability theory one may consult [3] and [6].)

The transformation C_0 which associates with $\{s_n\}$ the sequence $\{\pi_n\}$ can be thought of as a linear transformation on the space E^∞ of all scalar-valued sequences. By analogy with the finite-dimensional situation, one would expect C_0 to have a countably infinite family of eigenvalues, and indeed it does. For if m is an arbitrary nonnegative integer and $s^{(m)}$ is the sequence with entries $s_n^{(m)} = 0$ for $0 \leq n < m$,

$$s_{m+k}^{(m)} = \binom{m+k}{m} \text{ for } k = 0, 1, 2, \dots, \text{ then } C_0 s^{(m)} = \frac{1}{m+1} s^{(m)}.$$

Moreover, it is not difficult to prove that the numbers $(m+1)^{-1}$ are the only eigenvalues of C_0 and that each characteristic subspace is of dimension 1. We should note in passing that none of the eigenvectors for C_0 is a bounded sequence.

For each p , l^p is a linear subspace of E^∞ , and one may ask whether C_0 takes l^p into itself. That it does is a consequence of **Hardy's inequality** for sums, which asserts that

$$\sum |\pi_n|^p \leq q^p \sum |s_n|^p$$

and that the constant q^p cannot be replaced by any smaller number. (For a proof, see [7], p. 239.) In terms of operators we can state: the restriction of C_0 to l^p is an operator of norm q . Moreover, the restricted operator, which we shall continue to call C_0 , has no eigenvalues! Since the spectrum is the infinite-dimensional substitute for the set of eigenvalues, we are led to wonder: what is $\sigma(C_0, l^p)$?

Functions of a continuous variable can also be averaged. If f is measurable and is integrable over each subinterval $(0, x)$, then

$$(1) \quad F(x) = \frac{1}{x} \int_0^x f(t) dt$$

is defined and continuous. G. H. Hardy proved the following inequalities, which are analogous to his inequality for sums (see [7], p. 240):

$$\int_0^1 |F(x)|^p dx \leq q^p \int_0^1 |f(t)|^p dt,$$

$$\int_0^\infty |F(x)|^p dx \leq q^p \int_0^\infty |f(t)|^p dt,$$

and once again the constants are best possible. Thus there are operators C_1 on $L^p(0, 1)$ (the finite-range Cesàro integral operator) and C_∞ on $L^p(0, \infty)$ (the infinite-range Cesàro integral operator) which associate with each f in the given function space the function F given by (1). According to the inequalities, $\|C_1\| = q$ and $\|C_\infty\| = q$. Again the questions of spectra arise.

The eigenvalue problems are easily disposed of. Continuity of F implies, by the fundamental theorem of calculus, that any eigenvector must be continuously differentiable and must be a solution of the Euler differential equation $\lambda(xy)' = y$. (The possibility that 0 might be an eigenvalue must be dealt with separately.) The solutions are the constant multiples of x^β where $\beta = \lambda^{-1} - 1$. None of these lie in $L^p(0, \infty)$ and so C_∞ has no eigenvalues. On the other hand, since x^β belongs to $L^p(0, 1)$ if and only if the real part of p^β exceeds -1 , the set of eigenvalues of C_1 is the entire open disk $\{\lambda: \operatorname{Re}(\lambda^{-1}) > q^{-1}\}$.

The spectral facts for the operators C_0, C_1, C_∞ can be summarized as follows.

THEOREM 1. (a) $\sigma(C_0, l^p) = \{\lambda: \operatorname{Re}(\lambda^{-1}) \geq q^{-1}\} \cup \{0\}$

$$= \{\lambda: |\lambda - q/2| \leq q/2\}.$$

(b) $\sigma(C_1, L^p(0, 1)) = \sigma(C_0, l^p)$.

(c) $\sigma(C_\infty, L^p(0, \infty)) = \{\lambda: \operatorname{Re}(\lambda^{-1}) = q^{-1}\} \cup \{0\}$.

In order to prove (c), one exhibits specific formulas for the inverse of $\lambda I - C_\infty$ for λ exterior to the circle and for λ interior to the circle. For example, if $|\lambda - q/2| > q/2$, then $\lambda^{-1}I + \lambda^{-2}P_\lambda$ is the inverse of $\lambda I - C_\infty$, where

$$(P_\lambda f)(x) = \int_0^1 f(xt)t^{-1/\lambda} dt.$$

One then shows that no point λ_0 on the circle belongs to the resolvent set because $\|P_\lambda\|$ diverges to infinity as λ approaches λ_0 . The arguments for (a) and (b) are symmetric. Just as the interior of the disk $|\lambda - q/2| \leq q/2$ consists of eigenvalues for C_1 , the same open disk also consists of eigenvalues for C_0^* ; so both spectra contain the closed disk. On the other hand, if λ is exterior to the disk, the restriction, or rather the compression, of $\lambda^{-1}I + \lambda^{-2}P_\lambda$ to $L^p(0, 1)$ inverts $\lambda I - C_1$ and a sequence-to-sequence analogue of that operator inverts $\lambda I - C_0$. (For proofs see [1] and [11]. Note that C_0 acting on l^p has no eigenvalues, yet its adjoint has uncountably many eigenvalues—certainly not what one would naively guess about so simple an operator.)

Since $q = p/(p-1)$, the influence of the space l^p on the spectrum of C_0 is clearly seen, but one may ask what the disk in (a) has to do with the numbers $1/(n+1)$ which enter in the definition of C_0 . We shall find the answer below.

3. Hausdorff operators. Consider an arbitrary complex-valued measurable function k on the unit interval which satisfies the following integrability condition

$$(2) \quad \int_0^1 t^{-1/p} |k(t)| dt < \infty.$$

Associated with k are three linear operators which we shall call the Hausdorff operators determined by k . (That the transformations are indeed operators on the indicated spaces was proved by Hardy; see [6, Chapter XI].)

The **discrete operator** $S_0 = S_0(k)$ is defined by the formula

$$(3) \quad (S_0 s)_n = \sum_{m=0}^n \binom{n}{m} k_{n,m} s_m \quad (s \text{ in } l^p), \text{ where}$$

$$(4) \quad k_{n,m} = \int_0^1 (1-t)^{n-m} t^m k(t) dt.$$

The **finite-range operator** $S_1 = S_1(k)$ is defined by

$$(5) \quad (S_1 f)(x) = \int_0^1 f(xt) k(t) dt \quad (0 < x < 1, f \text{ in } L^p(0, 1)),$$

while the **infinite-range operator** is given by

$$(6) \quad (S_\infty f)(x) = \int_0^1 f(xt) k(t) dt \quad (0 < x < \infty, f \text{ in } L^p(0, \infty)).$$

Note that the kernel $k_1(t) \equiv 1$ yields the Cesàro operators C_0, C_1, C_∞ . The inverting operator P_λ mentioned in the proof of Theorem 1 corresponds to $k(t) = t^{-1/\lambda}$. Various choices of the kernel k yield other sequence-to-sequence operators and integral transforms which are well known to workers in summability theory. Let us list some of these here:

Cesàro means of order α	$k(t) = \alpha(1-t)^{\alpha-1}$
Hölder means of order α	$k(t) = \frac{1}{\Gamma(\alpha)} (\log t^{-1})^{\alpha-1}$
Gamma means Γ_a^α	$k(t) = \frac{a^\alpha}{\Gamma(\alpha)} t^{a-1} (\log t^{-1})^{\alpha-1}$
generalized Cesàro means $C_{a,\alpha}$	$k(t) = \frac{\Gamma(a+\alpha)}{\Gamma(a)\Gamma(\alpha)} t^{a-1} (1-t)^{\alpha-1};$

note in particular that Γ_a^1 corresponds to the kernel $k_a(t) = at^{a-1}$. (The corresponding integral operators take the form

$$F_a(x) = \frac{a}{x^a} \int_0^x u^{a-1} f(u) du.)$$

If the parameters are assumed to satisfy the conditions $a > 1/p$, $\alpha > 0$, then the integrability condition (2) is satisfied, and indeed, in each of the examples a strengthened condition prevails:

$$(2^*) \quad \int_0^1 t^{-\gamma-1/p} |k(t)| dt < \infty, \text{ for some } \gamma > 0.$$

Moreover, as one would expect from an averaging process, each summability kernel in the list is nonnegative and has total integral 1. (One may replace the absolutely continuous signed measures $k(t)dt$ by more general measures, but we shall not discuss the generalizations.)

If we forget the integrability conditions for the moment and consider any discrete Hausdorff transformation defined on E^∞ by (3) and (4), where k is merely integrable over $(0, 1)$, we find a curious result: The eigenvectors of $S_0(k)$ are the same as those for C_0 ; indeed, $S_0 s^{(m)} = \mu_m s^{(m)}$ for each nonnegative integer m , where

$$\mu_m = \int_0^1 t^m k(t) dt$$

is the m th moment of k . This leads one to jump to the conclusion that the spectral facts for S_0, S_1 , and S_∞ should be the same, in some sense, as those for C_0, C_1 , and C_∞ . And in a strong sense this is in fact so.

THEOREM 2. (a) S_0 has no eigenvectors in l^p , but if z is any scalar such that $\operatorname{Re}(z) > -1/p$, then the sequence f_z defined by

$$(1 - w)^z = \sum f_z(n) w^n$$

belongs to l^q and is an eigenvector of S_0^* . The corresponding eigenvalue is $K(z) = \int_0^1 t^z k(t) dt$. The spectrum of S_0 consists of 0 together with the range of $K(z)$ on the half-plane $\operatorname{Re}(z) \geq -1/p$.

(b) If $\operatorname{Re}(z) > -1/p$, then the function g_z defined by $g_z(t) = t^z$ belongs to $L^p(0, 1)$ and is an eigenvector of S_1 with corresponding eigenvalue $K(z)$. The spectrum of S_1 is the same as the spectrum of S_0 .

(c) If k satisfies condition (2^*) , then the spectrum of S_∞ is the union of $\{0\}$ with the range of $K(z)$ on the line $\operatorname{Re}(z) = -1/p$.

It is a consequence of the theorem that the spectrum of a Hausdorff operator is always a connected set containing the origin. The analytic geometry of the spectrum can be very intricate, as the reader may determine by calculating the moment function K for each of the examples.

Since $K_1(z) = \int_0^1 t^z k_1(t) dt = (z + 1)^{-1}$, Theorem 1 follows at once from Theorem 2. Moreover, the numbers $1/(n + 1)$ are just the moments $K_1(n)$. In general, the values of the moment function K , which the reader may perhaps recognize as a kind of Mellin transform of k on the half-plane or line with infinity adjoined, form the

spectrum while its values at the nonnegative integers are the diagonal entries $\mu_m = k_{m,m}$ in the infinite matrix representation of the discrete operator S_0 .

The facts about eigenvalues stated in Theorem 2 can be established by direct computation, once it is realized what they should be. The remaining assertions can be proved using facts about certain Banach algebras and groups of operators. These details will be published separately in [10]. We note also that D. W. Boyd has recently shown that conclusion (c) follows from the less stringent condition (2). See his forthcoming article entitled “Spectra of Convolution Operators.”)

4. Hilbert space considerations. Current interest in the Cesàro operators is due principally to the article [2] of Brown, Halmos, and Shields. These authors restrict themselves to the case where $p = 2$. Since l^2 , $L^2(0, 1)$, and $L^2(0, \infty)$ derive their norms from inner products, one might expect that a more delicate structural theory would prevail for the Hausdorff transformations viewed as operators on these Hilbert spaces. To an extent the expectations are justified.

Indeed, in [2] the following remarkable theorems are proved:

- (i) $I - C_1^*$ is a simple unilateral shift operator on $L^2(0, 1)$;
- (ii) $I - C_\infty^*$ is a simple bilateral shift operator on $L^2(0, \infty)$.

(Shifts are defined as follows. If X is a separable Hilbert space, an operator S on X is a simple unilateral shift provided that there is a maximal orthonormal sequence $\{e_n: n = 0, 1, 2, \dots\}$ in X such that $Se_n = e_{n+1}$ for every n . An operator T on X is a simple bilateral shift provided that there exists a maximal orthonormal bi-sequence $\{e_n: n = 0, \pm 1, \pm 2, \dots\}$ such that $Te_n = e_{n+1}$ for every n . The story of the shift operators is elegantly told in [4] and [5].)

The discrete operator cannot satisfy a condition analogous to (i) or (ii) since its adjoint has eigenvalues and shifts have none. The operator itself has some resemblance to a shift, but it is not one. Specifically, the following are true:

(iii) $I - C_0$ is unitarily equivalent to the operation $f(z) \rightarrow zf(z)$ on the closed subspace of $L^2(\beta)$ spanned by the polynomials, where β is a certain probability measure on the unit disk $|z| \leq 1$;

(iv) $I - C_0$ is not similar to any weighted shift. (We recall that operators A and B are similar if $A = PBP^{-1}$ for some invertible operator P ; if P can be chosen to be unitary, then A and B are unitarily equivalent. A weighted shift has the form SD where S is a shift and D is an operator which multiplies each e_n by a scalar β_n . See [8], [9].)

Theorems (i) and (ii) have been extended to the Gamma methods of order 1 (see [12]). The results read:

- (v) $(I - (2 - a^{-1})S_1(k_a))^*$ is a simple unilateral shift on $L^2(0, 1)$;
- (vi) $(I - (2 - a^{-1})S_\infty(k_a))^*$ is a simple bilateral shift on $L^2(0, \infty)$.

But one can generalize no further (see [10]):

(vii) If $k \geq 0$, $\int_0^1 k(t) dt = 1$, and $\int_0^1 t^{-1/2} k(t) dt < \infty$, then if for some $c > 0$, either $I - cS_1(k)^*$ or $I - cS_\infty(k)^*$ is a shift operator, then $k = k_a$ for some $a > \frac{1}{2}$.

Undoubtedly the summability kernels with $\alpha \neq 1$ will also turn out to be classifiable according to structural properties of the associated operators.

Some additional properties of Hausdorff operators on Hilbert spaces are known (for instance, S_∞ is always a normal operator and can be represented as the adjoint of a convolution operator on $L^2(-\infty, +\infty)$ generated by an integrable function which vanishes outside $(0, \infty)$) and more remain to be discovered. We hope that this article has convinced the reader that the algebraic and the classical points of view in analysis can each provide the other with a lively stimulus and a source of problems for study, and that "classical" need not mean "obsolete," even in the era of "now!"

References

1. D. W. Boyd, The spectrum of the Cesàro operator, *Acta Sci. Math.*, Szeged, 29 (1968) 31–34.
2. A. Brown, P. R. Halmos, and A. L. Shields, Cesàro operators, *Acta Sci. Math.*, Szeged, 26 (1965) 125–137.
3. R. G. Cooke, *Infinite Matrices and Sequence Spaces*, Macmillan, London, 1950.
4. P. R. Halmos, Shifts on Hilbert spaces, *J. Reine Angew. Math.*, 208 (1961) 102–112.
5. ———, *A Hilbert Space Problem Book*, Van Nostrand, Princeton, N. J., 1967.
6. G. H. Hardy, *Divergent Series*, Clarendon Press, Oxford, 1949.
7. G. H. Hardy, J. E. Littlewood, and G. Polya, *Inequalities*, reprinted second edition, Cambridge University Press, 1967.
8. T. L. Kriete and D. Trutt, The Cesàro operator in l^2 is subnormal, *Amer. J. Math.*, 93 (1971) 215–225.
9. ——— and ———, On the Cesàro operator, to appear.
10. G. Leibowitz, A convolution approach to Hausdorff integral operators, to appear.
11. B. E. Rhoades, Spectra of some Hausdorff operators, *Acta Sci. Math.*, Szeged, 32 (1971) 91–100.
12. N. K. Sharma, article to appear in *Acta Sci. Math.*, Szeged.
13. G. F. Simmons, *Introduction to Topology and Modern Analysis*, McGraw-Hill, New York, 1963.

A. A. ALBERT

D. ZELINSKY, Northwestern University

Abraham Adrian Albert died on June 6, 1972. The world lost a renowned mathematician, a vigorous force for the advancement of mathematics, and a very warm and understanding human being. From his birth to his death, he was associated with Chicago. As an inveterate traveller, he left that city often, for far parts of the world, but he always returned. He was born in Chicago on November 9, 1905, he went to school in Chicago (except for two years when his family moved to Iron Mountain, Michigan), he did all his undergraduate and graduate work at the Uni-

versity of Chicago. After receiving his Ph.D., he left for three years at Princeton and at Columbia Universities, then returned to the University of Chicago where he was a faculty member until the end of his life. With this as his base, he worked in many mathematical centers at various times in his career: The Institute for Advanced Study in Princeton (1933–34), Universities of Brazil and Buenos Aires (1947), University of Southern California (1950), Yale University (1956–1957), University of California at Los Angeles (1958). He operated in Washington in many capacities, and in the International Mathematical Union. His most recent official trip was a visit to the USSR in 1971 as a guest of the Soviet Academy.

To his friends Professor Albert was known as Adrian. Many mathematicians referred to him affectionately as A^3 . He was the son of a Jewish family that came to America from England. His father insisted on a Jewish but not very religious training. Albert distinguished himself early in his schools (Herzl and Marshall) on the West Side of Chicago, where the intellectual competition from the other budding scholars was keen. He spent four years earning his Bachelor's degree at the University of Chicago, but one year later he had his Master's degree, and a year after that, his Ph. D. In 1928, at age 22, his Ph. D. dissertation already stamped him as one of the outstanding algebraists of his day.

Those were the days when the mathematical leaders at the University of Chicago were L. E. Dickson in algebra and E. H. Moore in general topology. Dickson was Albert's thesis advisor and is the one mainly responsible for steering Albert into the subject of algebras over fields, which is the subject that primarily concerned him throughout his career.

He was one of the early National Research Council Fellows (1928–29). This fellowship was the forerunner of the modern NSF Postdoctoral Fellowships (which unfortunately were discontinued recently) and has been held by some of the most famous American mathematicians.

The precocity continued. At the age 35, Albert was promoted to a full professorship at the University of Chicago (at that time it was virtually unheard of to hold such a position before the age of 40). Two years later he was elected to membership in the National Academy of Sciences, a 37-year old academician.

The list of other honors heaped on him, and of honorific duties he was asked to perform would run to more pages than this article. We mention just a sample: chairmanship (1958–1962) and deanship (1962–1971) at the University of Chicago, presidency of the American Mathematical Society (1965–66), trusteeship of the institute for advanced Study (1969–72) chairmanship of the International Mathematical Union's organizing committee for the 1970 Congress in Nice, membership in the Brazilian and Argentine Academies of Sciences, several editorships, the Cole Prize in Algebra (1939), and three honorary degrees. He seemed to collect these honors with enthusiasm, and executed the duties with vigor.

Although Albert worked on matrix theory, on quadratic forms, and other aspects of algebra, there is no question that his central interest was always the study

of finite dimensional algebras over a field. In the old days, they were called hypercomplex systems. They are finite dimensional vector spaces with a multiplication that associates to every two vectors in the space another vector, the product. A suggestive example is the four-dimensional algebra of quaternions over the field of real numbers. The classical Wedderburn theorems essentially reduce the study of associative algebras over a field to the classification of the division algebras (like the algebra of quaternions, for example). Over any field F , a four-dimensional division algebra with center F must be an algebra of "generalized quaternions" whose multiplication rules are much like the ordinary quaternions: a basis $1, i, j, ij$ with $ij = ji$ and $i^2 = \alpha$ and $j^2 = \beta$ elements of F , which have no square roots in F (but are not necessarily -1). If one wants to generalize to dimensions higher than 4 there are two candidates: the cyclic algebras and the still more general crossed product algebras. (A theorem asserts that, in any case, the dimension of any central division algebra is a perfect square.) Wedderburn had already proved that central division algebras of dimension 9 are all cyclic algebras. In Albert's dissertation (1928) he proved that central division algebras of dimension 16 are not necessarily cyclic algebras, but are always crossed products. Although Albert's theorem raised the obvious question about algebras of dimension 25, 36, etc., his result has stood without essential improvement or embellishment (though not for lack of trying) until some nice, complementary, but still not definitive results of Amitsur and others in 1971.

This study put the young Albert in the center of what was to be one of the major breakthroughs in the theory of algebras: the determination of all central division algebras over the special field of rational numbers, or more generally over any algebraic number field. In this case, it turns out that they are all cyclic algebras—this is the famous Hasse-Brauer-Noether Theorem (1931). An interesting article by Hasse and Albert in the Transactions of the American Mathematical Society (1932) traces the history of this theorem and relates the story of Albert's near miss. On the basis of his results on algebras and some results announced by Hasse, Albert published some theorems that nearly proved the big theorem, and he wrote Hasse about it. Somehow the communication was bad, and the Brauer-Hasse-Noether manuscript was submitted for publication without mention of Albert's independent contributions. The 1932 Transactions article shows that in fact the big theorem follows from Albert's results in just a few lines.

Albert was hurt and disappointed by this incident. But the depth of that hurt could not compare with his feelings about the subsequent Nazi scourge which caused some important German mathematicians to begin distinguishing between "Aryan" and "Semitic" mathematics, and which resulted in the exodus of so many German scientists, Jews and non-Jews alike, including both Richard Brauer and Emmy Noether.

Albert was invited to be a member of the Institute for Advanced Study in Princeton during its opening year in 1933–34. (Another distinguished member, who arrived that year and remained on a permanent basis was Albert Einstein.) This contact

with Princeton was profitable for Albert. His associations with Lefschetz in particular resulted in one of Albert's mathematical accomplishments that he always regarded with greatest pleasure, and for which he later won the American Mathematical Society's Cole Prize in algebra. Already in 1929, Lefschetz had interested Albert in a major unsolved problem in the theory of algebraic functions, Riemann surfaces and Abelian varieties. In a series of papers (1929-1934) Albert produced a definitive solution. What was required was a classification of the algebraic correspondences of a Riemann surface (automorphisms of a complex curve). This had been reduced to the problem of finding the matrices that commute with a certain "Riemann matrix" of periods of basic Abelian integrals on the Riemann surface. These commuting matrices form an algebra, and in the basic cases, a central simple algebra over the rational number field. This version of the problem was right in the center of Albert's special expertise, and he demolished it.

Later, he attacked the problem of general nonassociative algebras that are finite dimensional over a field. Almost single-handed he influenced a large number of young mathematicians to break this seemingly unpromising ground. Special algebras had been studied that were not associative but which obeyed axioms substituting for the associative law (Lie algebras, Jordan algebras, alternative algebras). Results like the Wedderburn theorems had been proved for some of them; in fact, the results for Lie algebras over the complex number field were proved by E. Cartan before Wedderburn obtained his corresponding theorems for the associative algebras. But Albert had the idea of using associative algebra theory to prove analogs of the Wedderburn theorems for quite arbitrary nonassociative algebras (even a nonassociative algebra has an associative "regular representation" algebra). It is a sign of his genius that he was actually able to develop a reasonable theory and also significantly influence the theory and applications of the special algebras we mentioned.

Albert's style of algebra was almost inimitable. He had a diabolical facility with manipulation of identities—an enterprise in which most mathematicians founder, never being able to see the forest for the trees. Somehow, Albert could see through mazes of symbols to the inner workings of all those polynomials in several variables or multiplication tables of complicated algebras.

Mathematics was Albert's great enthusiasm. It was impossible to associate with him for any length of time without feeling the vigor with which he pursued his theorems. He was always willing to talk about his latest mathematical exploits. When his son, Alan, was still very young, Albert insisted on explaining even to him his latest theorems, patiently describing the necessary ingredients to the intrigued schoolboy who had not yet formally seen any real mathematics.

Albert was a prolific author of textbooks, research treatises, and more than a hundred and thirty research papers, the last of which is due to appear soon.

A minor theme running through Albert's life was his fascination with cameras, radios and other gadgets. I have always thought that this streak was responsible for his activity in the Applied Mathematics Group at Northwestern University

during World War II and his association with the Rand Corporation (1951 and 1952), Southern California Applied Mathematics Project (1953–55; he was its chairman 1959–60) and the Institute for Defense Analysis (director, Communication Research Division 1961–62, trustee 1969–72). His principal contribution in these activities was to cryptanalysis and coding. He was also a lifelong aficionado of detective stories, which he devoured at an enormous rate.

To his friends, Professor Albert is remembered as a devoted family man. He married Frieda Davis in his second year of graduate study, and they shared a close relationship for all the subsequent forty-four years. They had two sons and a daughter and five grandchildren. (Tragically one son died of an illness at the age of 25.) Perhaps one should also count his 29 Ph. D. students whom he treated almost as members of his family.

He was always pleased to use his influence in Washington to improve the status of mathematicians in general, and he was willing to do the same for individual mathematicians whom he considered worthy. One of the more homey causes to which he lent the weight of his reputation was retaining an apartment building at the University of Chicago for visiting mathematics faculty and their families. There are families throughout the world that remember this little mathematical microcosm with pleasure.

Everyone who knew him will remember his vigorous but round, medium build, curly hair, and often boyish demeanor; but especially one must remember his great, pleased grin that he flashed to welcome news of new successes for any of his extended family anywhere in the world of mathematics.

MATHEMATICAL NOTES

EDITED BY ROBERT GILMER

Material for this Department should be sent to David Roselle, Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803.

FUNCTIONS SATISFYING A MEAN VALUE PROPERTY AT THEIR ZEROS

D. P. STANFORD, College of William and Mary

We ask whether a function on the plane whose average at its zeros is zero, and which assumes the value zero on an open set, must be zero everywhere.

Specifically, let R denote the real numbers and for p in R^2 and $r > 0$, let $D(p, r)$ denote the open disc of radius r centered at p .

DEFINITION: A function F from R^2 to R has *property Z* provided $\iint_{D(p,r)} F = 0$ for all $r > 0$ whenever $F(p) = 0$.

We now ask: For what open sets U and what classes C of functions can we assert that if F is in C , F has property Z , and $F \equiv 0$ on U , then $F \equiv 0$ on R^2 ?

The question arises in the study of functions satisfying the weighted average property of A. K. Bose [1]; more specifically in trying to determine whether a non-constant function u , satisfying the weighted average property with respect to a positive continuous weight function w , could have local extreme values. (If w is in class C' , the answer is no, since u would have to satisfy the differential equation given by Bose.) If $u(p) = 0$ is a local extreme of u , then the function uw would have property Z and would be zero on an open set containing p .

THEOREM. *If $F(x, y) = f(x)g(y)$ is continuous from R^2 to R , F has the property Z , and $F \equiv 0$ on a set U open in R^2 , then $F \equiv 0$ on R^2 .*

The proof of the theorem depends on the following lemma:

LEMMA. *Suppose $T > 0$, f and h are continuous in $[0, 2T]$, $h(t) > 0$ for $0 < t < 2T$, and $\int_0^t f(u)h(t-u)du = 0$ for $0 \leq t \leq 2T$. Then $f(t) = 0$ for $0 \leq t \leq T$.*

The lemma is easily proved as a variation on the "Proof of Titchmarsh's theorem in the case $f = g$ " on page 20 of [2].

Proof of the theorem: U contains a non-void open rectangle $(a, b) \times (c, d)$ on which $F \equiv 0$. If there is an x_0 in (a, b) with $f(x_0) \neq 0$, then

$$g(y) = \frac{F(x_0, y)}{f(x_0)} = 0$$

for all y in (c, d) . Thus either $f \equiv 0$ on (a, b) or $g \equiv 0$ on (c, d) . We assume $f \equiv 0$ on (a, b) , the other case being similar.

Let (a^1, b^1) be the largest open interval (possibly infinite) containing (a, b) on which $f \equiv 0$. We shall show that $b^1 = \infty$. If $b^1 < \infty$ we may, by a translation, suppose $b^1 = 0$, so that $a^1 < 0$. If $f \equiv 0$ or $g \equiv 0$ on R , the theorem follows. Thus we assume that $f \not\equiv 0$ and $g \not\equiv 0$. It follows that f and g are continuous. Let q be a number with $g(q) \neq 0$. We assume $g(q) > 0$, the other case being similar. Let $T > 0$ such that $g(t) > 0$ for $q - 2T < t < q + 2T$ and $4T < -a^1$. Let

$$h(x) = \begin{cases} \int_{\alpha}^{\beta} g(t)dt, & 0 \leq x \leq 4T, \\ 0, & x > 4T, \end{cases}$$

where $\alpha = q - \sqrt{4T^2 - (x-2T)^2}$ and $\beta = q + \sqrt{4T^2 - (x-2T)^2}$. Then h is continuous and, for $0 < x < 2T$, the interval of integration in the definition of $h(x)$ is contained in $[q - 2T, q + 2T]$, so $h(x) > 0$.

We suppose $0 \leq t \leq 2T$ and show that $\int_0^t f(u)h(t-u)du = 0$. Since $2T < -a^1$

we have $a^1 < t - 2T \leq 0$. Thus $f(t - 2T) = 0$, so $F(t - 2T, q) = 0$. Now

$$D((t - 2T, q), 2T) = \{(x, y): t - 4T < x < t \text{ and} \\ q - \sqrt{4T^2 - (t - x - 2T)^2} < y < q + \sqrt{4T^2 - (t - x - 2T)^2}\}.$$

Since $f(x) = 0$ for $a^1 < x < 0$ and since $a^1 < t - 4T$, we have

$$\begin{aligned} 0 &= \iint_{D((t-2T, q), 2T)} F \\ &= \int_0^t \left\{ \int_\gamma^\delta f(u)g(y) dy \right\} du \\ &= \int_0^t f(u)h(t-u) du, \end{aligned}$$

where $\gamma = q - \sqrt{4T^2 - (t - u - 2T)^2}$ and $\delta = q + \sqrt{4T^2 - (t - u - 2T)^2}$. Thus by the Lemma, $f(t) = 0$ for $0 \leq t \leq T$. That is, (a^1, T) is an open interval on which $f \equiv 0$ which is larger than (a^1, b^1) , contrary to our choice of (a^1, b^1) . Thus $b^1 = \infty$. Similarly, $a^1 = -\infty$, and $f \equiv 0$ on R . Thus $F \equiv 0$ on R^2 and the theorem is proved.

References

1. A. K. Bose, Functions satisfying a weighted average property, Trans. Amer. Math. Soc., 118 (1965) 472-487.
2. Jan Mikusinski, Operational Calculus, Pergamon Press, New York, 1959.

ON AN EXTENSION OF THE THEOREM OF HAUSDORFF-YOUNG

LIANG-SHIN HAHN, University of New Mexico

The Hausdorff-Young theorem [1, p. 98] asserts that if $1 \leq p \leq 2$ and q is the conjugate exponent, that is, $1/p + 1/q = 1$, then

$$\left[\sum_{k \in \mathbb{Z}} |\hat{f}(k)|^q \right]^{1/q} \leq \|f\|_p, \quad f \in L^p(T).$$

If $p = 1$, the left-hand side is, by definition, $\sup_{k \in \mathbb{Z}} |\hat{f}(k)|$. (For notations, see Katznelson [1].)

Recalling that the Taylor coefficients of functions which are analytic in the (open) unit disc and continuous on the closed unit disc are precisely the Fourier coefficients of the boundary functions, the dramatic breakdown of the Hausdorff-Young theorem when $p > 2$, as shown here, does not seem to have been formulated before.

Let $HC(D)$ be the Banach space of all functions analytic in the (open) unit disc and continuous on the closed unit disc, endowed with the uniform norm.

THEOREM. For each function $f(z) = \sum_{k=0}^{\infty} \hat{f}(k)z^k$ in $HC(D)$, except perhaps those of a meager subset of $HC(D)$, it is the case that

$$\sum_{k=0}^{\infty} |\hat{f}(k)|^q = \infty \text{ for every } q < 2.$$

Proof. Let E and E_n^q ($n = 1, 2, 3, \dots$) be the sets of all functions $f(z) = \sum_{k=0}^{\infty} \hat{f}(k)z^k$ in $HC(D)$ such that $\sum_{k=0}^{\infty} |\hat{f}(k)|^q < \infty$ for some $q < 2$ and $\sum_{k=0}^{\infty} |\hat{f}(k)|^q \leq n$, respectively. Since $E = \bigcup_{n=1}^{\infty} E_n^q$, where $\{q_n\}_{n=1}^{\infty}$ is a monotone increasing sequence of positive real numbers converging to 2, it is sufficient to show that each E_n^q is nowhere dense ($q < 2$). The proof is carried out in two stages:

1. The set E_n^q is closed. If $f_j \rightarrow f$ uniformly, then $\hat{f}_j(k) \rightarrow \hat{f}(k)$ for all k (uniformly) as $j \rightarrow \infty$, hence for every positive integer M ,

$$\sum_{k=0}^M |\hat{f}_j(k)|^q \rightarrow \sum_{k=0}^M |\hat{f}(k)|^q \quad (j \rightarrow \infty).$$

Thus, if $f_j \in E_n^q$ ($j = 1, 2, 3, \dots$), then $\sum_{k=0}^M |\hat{f}(k)|^q \leq n$, and $f \in E_n^q$.

2. Suppose $f \in HC(D)$. We want to show that given any $\varepsilon > 0$, there exists a function g in $HC(D)$, but not in E_n^q , such that $\|f - g\| < \varepsilon$. Since E_n^q is closed, this will establish that E_n^q is nowhere dense. Choose a polynomial g_1 such that $\|f - g_1\| < \varepsilon/2$. Existence of such a polynomial is guaranteed by the Fejér theorem (and the maximum modulus principle).

Next, we claim that given any $\eta > 0$ and $0 < \xi < 1$, there exists a polynomial h satisfying $\xi < |h(z)| < 1$ for all $|z| = 1$, and $\max_k |\hat{h}(k)| < \eta$.

Once the existence of such a function is established, then since

$$\sum_{k=0}^n |\hat{h}(k)|^q > \eta^{q-2} \sum_{k=0}^n |\hat{h}(k)|^2 = \eta^{q-2} \frac{1}{2\pi} \int_0^{2\pi} |h(e^{it})|^2 dt > \eta^{q-2} \xi^2, \quad (n = \deg h),$$

(where in the second step, we have used the Parseval equality for the polynomial h), we may let $g_2(z) = (\varepsilon/2)z^{2m}h(z)$, where $m = \deg g_1$, then $g = g_1 + g_2$ satisfies the condition prescribed above. Clearly, $g \in HC(D)$,

$$\|f - g\| \leq \|f - g_1\| + \|g_2\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

and

$$\begin{aligned} \sum_{k=0}^{\infty} |\hat{g}(k)|^q &= \sum_{k=0}^m |\hat{g}_1(k)|^q + \sum_{k=2m}^{2m+n} |\hat{g}_2(k)|^q \\ &= \sum_{k=0}^m |\hat{g}_1(k)|^q + \left(\frac{\varepsilon}{2}\right)^q \sum_{k=0}^n |\hat{h}(k)|^q \\ &> \sum_{k=0}^m |\hat{g}_1(k)|^q + \left(\frac{\varepsilon}{2}\right)^q \cdot \eta^{q-2} \cdot \xi^2. \end{aligned}$$

Thus, noting that $q - 2 < 0$, we may, by choosing η sufficiently small (and $\xi = \frac{1}{2}$, say), make the last term exceed n ; hence $g \notin E_n^q$.

It remains to show the existence of a polynomial h satisfying the required conditions. Choose M and δ , $0 < \delta < 1$, so that $(\frac{2}{3})^M < \eta$, $\xi < \delta^M < 1$. By taking a suitable partial sum $\phi_N(z)$ of the Taylor series expansion of

$$\phi(z) = \frac{z-a}{1-\bar{a}z} \quad \left(|a| = \frac{2}{3} \right),$$

one can get $\delta < |\phi_N(z)| < 1$ for $|z| = 1$. Consider

$$h(z) = \prod_{j=0}^{M-1} \phi_N(z^{\lambda^j}) \quad (\lambda = 2N).$$

Then $\xi < \delta^M < |h(z)| < 1$, for $|z| = 1$.

The Taylor coefficients of $h(z)$ are (either zero or) products of M Taylor coefficients of $\phi_N(z)$, but since the maximum modulus of the Taylor coefficients of $\phi(z)$ (hence of $\phi_N(z)$ also) is $\frac{2}{3}$, we have $|h(k)| \leq (\frac{2}{3})^M < \eta$ for all k . Thus $h(z)$ satisfies the required conditions and we are done.

The proof which has been presented also establishes the following:

COROLLARY. *In the Banach space $C(T)$ of all continuous functions on the circle T (with uniform norm), except perhaps those of a meager subset, each function has a sequence of Fourier coefficients that belongs to $l^q(\mathbb{Z})$ for no value of $q < 2$.*

Like many problems in analysis, this is a problem of comparison of norms. If a family of new "norms" are lower semi-continuous with respect to the original Banach space norm, then the existence of one "counterexample" implies that the collection of all "counterexamples" fills up the space to within a meager subset. We leave it as an exercise for the reader to investigate the boundary cases of the M. Riesz conjugate function theorem [1, p. 68] and the Bernstein theorem on absolute convergence of Fourier series [1, p. 32].

In conclusion we mention an open problem: Give an example (or show the existence) of a subset E of non-negative integers so that whenever $\sum_{k \in E} \hat{f}(k)z^k$ is the Taylor series of a function f in $HC(D)$, then

$$\sum_{k \in E} |\hat{f}(k)|^p < \infty \quad \text{for all } p > 1,$$

but there is a Taylor series $\sum_{k \in E} \hat{g}(k)z^k$ of such a function g with $\sum_{k \in E} |\hat{g}(k)| = \infty$.

Acknowledgement. The author wishes to thank Professor B. Epstein for his help in preparation of the manuscript.

Reference

1. Y. Katznelson, *An Introduction to Harmonic Analysis*, Wiley, New York, 1968.

A CHARACTERIZATION OF THE $n \times n$ MATRICES OVER A FINITE FIELD

J. V. BRAWLEY, Clemson University, and L. CARLITZ, Duke University

If R is a commutative ring with identity, then it is known [4, p. 507] that every function $f: R \rightarrow R$ is representable as a polynomial if and only if R is a finite field. The theorems below are a generalization of that result to arbitrary rings with or without an identity.

Consider an expression in an indeterminate x of one of the following types:

$$(1) \quad f(x) = \begin{cases} a_1 x^{n_1} a_2 x^{n_2} \cdots a_k x^{n_k} a_{k+1}, \\ a_1 x^{n_1} a_2 x^{n_2} \cdots a_k x^{n_k}, \\ x^{n_1} a_1 x^{n_2} a_2 \cdots x^{n_k} a_k, \\ x^{n_1} a_1 x^{n_2} a_2 \cdots a_k x^{n_{k+1}}, \\ x^{n_1}, \end{cases}$$

where the elements a_i are in R , the integers n_i are positive, and $k \geq 1$ is finite but arbitrary. In each of these expressions it is meaningful to speak of the function $f: R \rightarrow R$ defined from $f(x)$ by substitution. Likewise if $p(x)$ is a form of the type

$$(2) \quad p(x) = a_0 + f_1(x) + \cdots + f_t(x),$$

where each $f_i(x)$ is of the type (1), $a_0 \in R$, and $t \geq 0$, it is clear that substitution in (2) defines a function p from R to R . An expression of the form (2) will be called a **polynomial over R** (this definition of polynomial is more general than that which is often considered, e.g. [3] p. 99), and $P(R)$ will denote all those functions in R^R which can be represented by polynomials (upon substitution). We seek necessary and sufficient conditions on R in order that every function from R to R be representable as a polynomial; that is, that $R^R = P(R)$.

LEMMA 1. If $R^R = P(R)$, then R is finite.

Proof. Suppose, to the contrary, that $\text{card } R = |R| = \infty$. Then the number of expressions of the form (2) has cardinality $|R|$ so that $|R| = |P(R)| = |R^R| = |R|^{|R|}$. This is a contradiction as $|R|^{|R|} > |R|$.

LEMMA 2. If $R^R = P(R)$, then the only ideals of R are (0) and R .

Proof. Suppose there exists an ideal I in R with $0 \subsetneq I \subsetneq R$. Let $\mu: R \rightarrow R/I$ be the natural homomorphism, $\mu(r) = \bar{r} = r + I$, and select $a \in I$, $a \neq 0$ and $b \notin I$. Since $R^R = P(R)$ there exists a polynomial of the form (2) representing the function

$$p(r) = \begin{cases} b; & r = a \\ 0; & r \neq a. \end{cases}$$

Let $\bar{p}(x)$ be the polynomial over R/I obtained from (2) by replacing each $a_i \in R$ by \bar{a}_i . Clearly $\bar{p}(\bar{r}) = \overline{p(r)}$ as μ is a homomorphism. Now $\bar{p}(\bar{0}) = \bar{p}(\bar{a})$ as $\bar{a} = \bar{0}$. But $\bar{p}(\bar{0}) = \overline{p(0)} = \bar{0}$ and $\bar{p}(\bar{a}) = \overline{p(a)} = \bar{b} \neq \bar{0}$, which is a contradiction.

We can now prove

THEOREM 1. *Let R be a ring. Then $R^R = P(R)$ if and only if R is either the trivial ring of order 1 or 2 ($ab = 0 \forall a, b \in R$) or for some n and some finite field F , $R = (F)_n$, the $n \times n$ matrices over F .*

Proof. Suppose that $R^R = P(R)$. From the above lemmas, R is finite and has only trivial ideals. Assume first that R is a trivial ring. Then $(R, +)$ has no proper subgroups as every subgroup is an ideal. This means that R has order 1 or p , p prime [3, p. 42, Ex. 3.29]. Because R is a trivial ring it is clear that polynomials of the form $p(x) = a_0 + a_1x$, where $a_0, a_1 \in \{0, 1, \dots, p-1\}$ must represent all the functions in R^R . If $|R| = p$, the number of such polynomials is p^2 so that $p^2 = p^p$ or $p = 2$. If R is not the trivial ring then since R is a finite simple ring we have by the Wedderburn-Artin theorem [see 2, p. 48] and the Wedderburn theorem on finite division rings [see 2, p. 70] that $R = (F)_n$ for some finite field F and integer n .

Conversely, if R is a trivial ring of order 1 or 2, it is easy to see that every function in R^R is a polynomial; hence, assume $R = (F)_n$ for some $n \geq 1$ and finite field F . To complete the proof we shall use the following lemma:

LEMMA 3. *Let R be a finite ring with identity. Then $R^R = P(R)$ iff the function*

$$(3) \quad p(r) = \begin{cases} 1; & r = 0 \\ 0; & r \neq 0 \end{cases}$$

is representable by a polynomial $p(x)$.

Proof of Lemma 3. If $R^R = P(R)$, then of course, p of (3) is representable by a polynomial. On the other hand, if $p(x)$ represents p and if f is an arbitrary function from R to R , then

$$f(x) = \sum_{r \in R} f(r)p(x-r)$$

is a polynomial representation of f .

Returning to the proof of the theorem we need only to show that if $R = (F)_n$ then the function (3) is representable by a polynomial of the type

$$f(X) = A_0 + \sum_{i=1}^m f_i(X),$$

where the polynomials $f_i(X)$ are of the type $A_1XA_2 \cdots A_kXA_{k+1}$ ($k \geq 1$), with $A_i \in (F)_n$ and with $X = (x_{ij})$ denoting an $n \times n$ matrix in indeterminates x_{ij} . Note since $(F)_n$ has an identity (denoted by I) only one of the forms of (1) is necessary.

By a result of Dickson [1, p. 124], there exists a polynomial $f(\bar{x})$ in the n^2 var-

ables $\bar{x} = (x_{11}, \dots, x_{1n}, x_{21}, \dots, x_{2n}, \dots, x_{n1}, \dots, x_{nn})$ which upon substitution of values from F , takes on the values 0 for $\bar{x} = 0$ and 1 for $\bar{x} \neq 0$. Denoting by E_{ij} the $n \times n$ elementary matrix over F with a 1 in position (i, j) and zeros elsewhere, it is easily seen that $E_{ti} X E_{jt} = x_{ij} E_{tt}$. Setting $X_{ij}^{(t)} = E_{ti} X E_{jt}$ and $X^{(t)} = (X_{11}^{(t)}, X_{12}^{(t)}, \dots, X_{nn}^{(t)})$, we find that $f(\bar{x}^{(t)}) = f_t(X)$ is a polynomial in X and $f_t(X) = f(\bar{x}) E_{tt}$. Thus

$$p(X) = I - \sum_{t=1}^n f_t(X) = I - \text{diag}(f(\bar{x}), f(\bar{x}), \dots, f(\bar{x}))$$

is a polynomial in X taking on the value I for $X = 0$ and 0 for all $X \neq 0$. Thus by Lemma 3 the proof is complete.

Since (F) is commutative iff $n = 1$ we have the following corollary which includes the result mentioned in the first paragraph.

COROLLARY. *Let R be a commutative ring. Then $R^R = P(R)$ iff either R is a trivial ring of order 1 or 2, or R is a finite field.*

By restricting the notion of polynomials we obtain as a final result:

THEOREM 2. *Let R be a ring with 1 and let $L[x]$ and $R[x]$ denote respectively the left and right polynomials over R ; that is, $L[x] = \{a_0 + xa_1 + \dots + x^n a_n; a_i \in R\}$, $R[x] = \{a_0 + a_1 x + \dots + a_n x^n; a_i \in R\}$. Let $P_L(R)$ and $P_R(R)$ denote those functions in R^R representable by left and right polynomials, respectively. Then $P_L(R) = R^R$ iff $P_R(R) = R^R$ iff R is a finite field.*

Proof. Suppose $P_L(R) = R^R$. Then R is finite by an argument similar to that of Lemma 1. Let $b \in R$ be an arbitrary nonzero element and let $p(x) = a_0 + xa_1 + \dots + x^n a_n$ be the left polynomial representing the function $p: R \rightarrow R$ defined by $p(b) = 1$, $p(r) = 0$ for $r \neq b$. Since $p(0) = 0$ it follows that $a_0 = 0$ and since $p(b) = 1 = b(a_1 + \dots + b^{n-1} a_n)$ it follows that b has a right inverse, i.e., $R - \{0\}$ is a group. Thus R is a finite division ring which by Wedderburn's theorem is a field. A similar proof holds if $P_R(R) = R^R$.

If R is a finite field it is well known that $R^R = P_L(R) = P_R(R)$ and the proof is complete.

L. Carlitz's work was supported in part by NSF grant GP-17031.

References

1. L. E. Dickson, General theory of modular invariants, Trans. Amer. Math. Soc., 10 (1909) 123-158.
2. I. N. Herstein, Noncommutative Rings, Carus Mathematical Monograph, No. 15, M.A.A. (1968).
3. N. Jacobson, Lectures in Abstract Algebra, Vol. 1, Van Nostrand, Princeton, N.J., 1953.
4. L. Rédei, Algebra, Vol. 1, Pergamon Press, Oxford, 1967.
5. J. J. Rotman, The Theory of Groups: An Introduction, Allyn and Bacon, Boston, 1965.

ANOTHER PROOF OF BERNSTEIN'S THEOREM

P. J. O'HARA, Florida Technological University

If P is a complex polynomial put

$$\|P\|_R = \max_{|z|=R} |P(z)|.$$

A well-known theorem states that if P is of degree n then

$$\|P'\|_R \leq \frac{n}{R} \|P\|_R.$$

In the literature (see [1] or [2]) this result is usually derived from the equivalent theorem by S. N. Bernstein [3] for trigonometric polynomials on the real line. The proof for the trigonometric case usually involves some analysis such as Rolle's theorem. In this note an apparently new proof is given which in essence involves no analysis. The proof depends on the following lemma which is of interest in itself.

LEMMA. If P is any complex polynomial of degree $\leq n$ and z_1, \dots, z_n are the zeros of $z^n + 1$, then for every complex number t ,

$$tP'(t) = \frac{n}{2}P(t) + \frac{1}{n} \sum_1^n P(tz_k) \frac{2z_k}{(z_k - 1)^2}.$$

Proof. For each complex number t define the function Q_t by:

$$Q_t(z) = \frac{P(tz) - P(t)}{z - 1}.$$

Q_t is a polynomial of degree $\leq n - 1$, and therefore by using the Lagrange interpolation formula [4] with z_1, \dots, z_n as interpolation nodes we can write:

$$Q_t(z) = \sum_1^n Q_t(z_k) \frac{z^n + 1}{n z_k^{n-1}(z - z_k)} = \frac{1}{n} \sum_1^n Q_t(z_k) \frac{z^n + 1}{z_k - z} z_k.$$

In obtaining this last equation we have used the fact that $z_k^{n-1} = -1/z_k$. Now since $Q_t(1) = tP'(t)$, we obtain the following identity in t :

$$\begin{aligned} tP'(t) &= \frac{1}{n} \sum_1^n Q_t(z_k) \frac{2z_k}{z_k - 1} = \frac{1}{n} \sum_1^n [P(tz_k) - P(t)] \frac{2z_k}{(z_k - 1)^2} \\ (1) \quad &= \frac{1}{n} \sum_1^n P(tz_k) \frac{2z_k}{(z_k - 1)^2} - \frac{P(t)}{n} \sum_1^n \frac{2z_k}{(z_k - 1)^2}. \end{aligned}$$

Applying (1) with $P(t) = t^n$ establishes that

$$(2) \quad \frac{1}{n} \sum_1^n \frac{2z_k}{(z_k - 1)^2} = -\frac{n}{2}.$$

Clearly (1) and (2) establish the lemma.

In order to prove Bernstein's theorem we first apply the lemma to obtain that if $|t| = R$ then

$$(3) \quad R |P'(t)| \leq \left[\frac{n}{2} + \frac{1}{n} \sum_1^n \left| \frac{2z_k}{(z_k - 1)^2} \right| \right] \|P\|_R.$$

Now if $|z| = 1$ and $z \neq 1$ then $2z/(z-1)^2$ is a negative real number. In fact it is easy to show that

$$\frac{2e^{i\theta}}{(e^{i\theta} - 1)^2} = -\frac{1}{2\sin^2 \theta/2}, \quad \theta \neq 0(2\pi).$$

Bernstein's theorem follows easily now from (2) and (3).

The lemma and other similarly derived identities can be used to obtain many other interesting results. For example the following theorem for which the author has no reference is an immediate consequence of the above discussion.

THEOREM. *If P is a complex polynomial of degree n , $P(t_0) = 0$, and $|t_0| = R$ then*

$$|P'(t_0)| \leq \frac{n}{2R} \|P\|_R.$$

References

1. I. P. Natanson, *Constructive Function Theory*, vol. 1, Ungar, New York, 1964, p. 90 (Translated from the Russian.)
2. G. G. Lorentz, *Approximation of Functions*, Holt, Rinehart, and Winston, New York, 1966, p. 39.
3. S. N. Bernstein, Sur l'ordre de la meilleure approximation des fonctions continues par des polynômes de degré donné, *Mémoires de l'Académie Royale de Belgique*, 2 (1912), vol. 4, pp. 1-104.
4. F. B. Hildebrand, *Introduction to Numerical Analysis*, McGraw-Hill, New York, 1956, p. 62.

ADDENDUM TO "ON THE DIFFEOMORPHISMS OF EUCLIDEAN SPACE"

W. B. GORDON, Naval Research Laboratory, Washington

A recent note [1] in this MONTHLY was concerned with the following theorem, which is a global version of the standard Inverse Function Theorem.

THEOREM. *Let M_1 and M_2 be connected, oriented manifolds of class C^1 , without boundary, and suppose that M_2 is simply connected. Then a C^1 map f from M_1 to M_2 is a diffeomorphism if and only if f is proper and the Jacobian of f never vanishes.*

In particular, a C^1 map f from R^N to R^N is one-one and onto if

- (i) $|x| \rightarrow \infty$ implies $|f(x)| \rightarrow \infty$, and
- (ii) *The Jacobian of f never vanishes.*

In this note I observed that although this theorem was known to Hadamard (1905), it does not appear to be "well-known" at the present time, and soon after the note went to press I learned that the theorem was rediscovered by R. S. Palais in the course of developing propositions of a more general character. See [2, p. 128–129].

Professor Palais also suggests the following simplifications to the proof: In my note I appealed to the standard results of degree theory to establish that f is onto; but this is not necessary, for one can easily show that a proper map between manifolds sends closed sets into closed sets. (For a more general statement of this fact see [3].) But the non-vanishing of the Jacobian implies that f is a local homeomorphism and therefore sends open sets into open sets. Hence $f(M_1)$ is both an open and closed (nonempty) subset of M_2 ; i.e., $f(M_1) = M_2$.

References

1. W. B. Gordon, On the diffeomorphisms of euclidean space, this MONTHLY, 79 (1972) 755–759.
2. R. S. Palais, Natural operations on differential forms, Trans. Amer. Math. Soc., 92 (1959) 125–141.
3. ———, When proper maps are closed, Proc. Amer. Math. Soc., 24 (1970) 835–836.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

HOW UNEXPECTED IS THE PRIME NUMBER THEOREM?

M. D. HIRSCHHORN, University of New South Wales, Australia

Let p_1, \dots, p_n be the first n primes. Let x be chosen randomly from among the integers greater than p_n . The probability that x is divisible by p_i is $1/p_i$, so the probability that x is divisible by none of p_1, \dots, p_n is

$$\left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_n}\right).$$

If we make the unjustified assumption that the property of being divisible by none of p_1, \dots, p_n is held randomly by numbers greater than p_n with probability $(1 - 1/p_1) \cdots (1 - 1/p_n)$, then the probability that

$$p_{n+1} - p_n = r$$

is

$$\left(1 - \left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_n}\right)\right)^{r-1} \left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_n}\right).$$

So the expected value of $p_{n+1} - p_n$ is

$$\begin{aligned} \sum_{r=1}^{\infty} r \left(1 - \left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_n}\right)\right)^{r-1} \left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_n}\right) \\ = 1 / \left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_n}\right). \end{aligned}$$

Accordingly, we define the series of **problimes** by

$$q_1 = 2, \quad q_{n+1} = q_n + 1 / \left(1 - \frac{1}{q_1}\right) \cdots \left(1 - \frac{1}{q_n}\right).$$

The interesting question that arises is, what is the asymptotic behavior of the q_n ? In particular, is it true that

$$q_n \sim n \log_e n?$$

If this *is* true, then perhaps the prime number theorem is a little less surprising, since it is a consequence of the prime number theorem that

$$p_n \sim n \log_e n.$$

I have been able to prove that

(1) q_n/n is unbounded,

(2) $q_n/n^{1+\varepsilon} \rightarrow 0$ as $n \rightarrow \infty$ for any fixed $\varepsilon > 0$,

but I have not been able to prove that q_n/n is eventually monotonic increasing.

It is clear that the problimes soon become non-integral. It may be more attractive to study the various integral sequences defined by

$$q_1 = 2, \quad q_{n+1} = q_n + F\left(1 / \left(1 - \frac{1}{q_1}\right) \cdots \left(1 - \frac{1}{q_n}\right)\right),$$

where

$F(x) = [x]$, the greatest integer not greater than x ,

$F(x) = \langle x \rangle$, the closest integer to x , or

$F(x) = \{x\}$, the least integer not less than x .

The first few terms of each of these sequences are given in the following table, together with the primes for comparison:

n	1	2	3	4	5	6	7	8	9	10	11
p_n	2	3	5	7	11	13	17	19	23	29	31
q_n^*	2.0	4.0	6.7	9.8	13.3	17.1	21.1	25.3	29.7	34.2	38.9
$q_n, []$	2	4	6	9	12	15	19	23	27	31	35
$q_n, \langle \rangle$	2	4	7	10	13	17	21	25	29	34	39
$q_n, \{ \}$	2	4	7	11	15	19	23	28	33	38	43

n	12	13	14	15	16	17	18	19	20	21
p_n	37	41	43	47	53	59	61	67	71	73
q_n	43.7	48.6	53.6	58.7	63.9	69.2	74.6	80.0	85.5	91.1
$q_n, []$	40	45	50	55	60	65	70	75	80	86
$q_n, \langle \rangle$	44	49	54	59	64	69	74	79	84	90
$q_n, \{ \}$	48	53	58	63	68	73	79	85	91	97

* to 1 decimal place

P. Erdős (written communication) believes that he can prove by Tauberian arguments that $(q_{n+1} - q_n) / \log_e n \rightarrow 1$, and hence that $q_n / n \log_e n \rightarrow 1$.

I am indebted to the referee and to Richard K. Guy for their helpful comments and suggestions.

CLASSROOM NOTES

EDITED BY ROBERT GILMER

Material for this Department should be sent to David Roselle, Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803.

THE INDECOMPOSABILITY OF THE DYADIC SOLENOID

S. B. NADLER, JR., University of North Carolina at Charlotte

In a beginning topology course students are sometimes confronted with the notion of "indecomposability" for continua (a continuum is *indecomposable* [4, p. 139] provided it cannot be written as the union of two proper subcontinua). There are proofs in the literature (see, for example, [3], [5], and material related

to [5]) that certain continua are indecomposable, but these proofs are "complicated" and, for the most part, inaccessible to a beginning student. Intuitive arguments for the indecomposability of certain continua, based on thinking of them as special nested intersections, are often dissatisfying to students because some of the details of such arguments are difficult to write down.

The purpose of this note is to give (in detail) a short and direct proof of the well-known and often-quoted result (see, for example, [4, p. 145]) that the dyadic solenoid is indecomposable. The proof is the simplest complete proof of which the author is aware, showing the indecomposability of any specific continuum. It uses only the very basic properties of inverse limits, a construction which appears in many standard texts ([2], [4], and others). These properties all appear in [1], where it is shown that they are a consequence of translating, to the framework of inverse limits, such standard results as (for example) the intersection of a decreasing sequence of metric continua is a metric continuum, etc. The crux of the proof is a pigeonhole-type argument.

Let X denote the inverse limit of the inverse sequence $\{C_i, f_i\}_{i=1}^{\infty}$ where C_i is the unit circle in the plane and $f_i: C_{i+1} \rightarrow C_i$ is given by $f_i(z) = z^2$ for each $i = 1, 2, \dots$ (X is commonly called the dyadic solenoid). In what follows we let $\pi_i: X \rightarrow C_i$ be the projection mapping (restricted to X).

THEOREM. X is indecomposable.

Proof. Let A and B be subcontinua of X such that $A \cup B = X$. It suffices to show that $A \subset B$ or $B \subset A$. Let i be a given natural number and suppose $\pi_i(A) \not\subset \pi_i(B)$ and $\pi_i(B) \not\subset \pi_i(A)$. Choose $a \in \pi_i(A) - \pi_i(B)$ and $b \in \pi_i(B) - \pi_i(A)$, let x and y be the two square roots of a , and let z and w be the two square roots of b . Since the bonding maps are all onto, the projection π_{i+1} maps X onto C_{i+1} (2.6 of [1]). Thus, since $A \cup B = X$,

$$x \in \pi_{i+1}(A) \cup \pi_{i+1}(B).$$

If $x \in \pi_{i+1}(B)$, then

$$a = f_i(x) \in f_i(\pi_{i+1}(B)) = \pi_i(B)$$

(2.1 of [1]), a contradiction. Therefore, $x \in \pi_{i+1}(A) - \pi_{i+1}(B)$ and the same is true of y . Similarly, z and w each belong to $\pi_{i+1}(B) - \pi_{i+1}(A)$. Since A is connected and x and y are diametrical members of $\pi_{i+1}(A)$, $\pi_{i+1}(A)$ must contain at least one of the two semicircles determined by x and y . But then z and w being diametrical implies z or w is in this semicircle, hence in $\pi_{i+1}(A)$, a contradiction. This proves

$$(1) \pi_i(A) \subset \pi_i(B) \quad \text{or} \quad (2) \pi_i(B) \subset \pi_i(A).$$

Thus, (1) holds for infinitely many natural numbers or (2) holds for infinitely many

natural numbers. This implies, by 2.1 of [1], that (1) holds for all natural numbers or (2) holds for all natural numbers. Now, since A is equal to the inverse limit of

$$\{\pi_i(A), f_i \mid \pi_{i+1}(A)\}_{i=1}^{\infty}$$

and B is equal to the inverse limit of

$$\{\pi_i(B), f_i \mid \pi_{i+1}(B)\}_{i=1}^{\infty}$$

(see 2.8 of [1]), it follows that $A \subset B$ or $B \subset A$.

References

1. C. E. Capel, Inverse limit spaces, *Duke Math. J.*, 21(1954) 233–245.
2. James Dugundji, *Topology*, Allen and Bacon, Boston, Mass., 1966.
3. A. van Heemert, Topologische Gruppen und unzerlegbare Kontinua, *Compositio Math.*, 5(1937) 319–326.
4. John G. Hocking and Gail S. Young, *Topology*, Addison-Wesley, Reading, Mass., 1961.
5. W. T. Ingram, Concerning nonplanar circle-like continua, *Canadian J. Math.*, 19 (1967) 242–250.

THE DIFFERENTIABILITY PROPERTIES OF TYPICAL FUNCTIONS IN $C[a, b]$

A. M. BRUCKNER, University of California, Santa Barbara

1. Introduction. Most students who complete an undergraduate program in mathematics encounter in some course an example of a continuous nowhere differentiable function. They may or may not be surprised to learn of the existence of such functions, but they are generally surprised that such functions represent the rule, rather than the exception. In saying that continuous nowhere differentiable functions represent the rule we mean that if $C[a, b]$ represents, as usual, the set of continuous real valued functions defined on the closed interval $[a, b]$ furnished with the metric of uniform convergence, then the subset of nowhere differentiable functions is residual in the complete metric space $C[a, b]$. This means that the complement of this subset is a set of the first category and is therefore negligible in the sense of category. In later courses, the student comes across other types of “pathological” continuous functions and learns that they, too, are typical, rather than exceptional. In any case, the student might complete his mathematical education with a feeling that typical continuous functions behave very irregularly with respect to differentiation. And he may wonder — “what is the typical continuous function like when it comes to differentiation properties?” The purpose of this note is to provide an answer to this question. We shall see that the typical continuous function is indeed very pathological. But, by taking a different perspective, we shall see that one can also come to the opposite conclusion.

When we say a “typical” continuous function has a certain property, we shall

mean that the set of continuous functions with this property is residual in $C[a, b]$. We mention that none of the material we present is new. It can all be found in the mathematical literature, but only parts of what we present have found their way into the textbooks.

2. Nowhere differentiable functions. The first example of a continuous nowhere differentiable function is widely assumed to be due to Weierstrass (about 1875), although there appears to be some evidence suggesting that Bolzano had constructed such a function considerably earlier. More recently, a number of authors have constructed simpler examples of such functions. See for example Mikolás [8] for a general method of constructing continuous nowhere differentiable functions. It wasn't until 1931 that the existence of such functions was proved by the use of the Baire Category Theorem (see Banach [1] and Mazurkiewicz [7]). Now to say that a function is nowhere differentiable is to say that at no point does it have a finite (two-sided) derivative. What happens if we allow derivatives to be infinite? And what happens if we allow derivatives to be only one-sided? An inspection of the standard category argument shows that one can, by perhaps modifying the arguments a bit, allow *either* of these relaxations in the definition of the derivative and still conclude that a typical continuous function is nowhere differentiable. But one can't allow both relaxations simultaneously without losing the result! Thus, we can say that a typical continuous function has at no point a finite or infinite two-sided derivative, nor a finite one-sided derivative, but does have an infinite one-sided derivative at each point of some nonempty set S . Actually, it was shown by Saks [10] that this set S is nondenumerable. So we can't use a category argument to determine whether or not there exists a continuous function which fails at every point to possess a one-sided finite or infinite derivative. The question of the existence of such functions remained unanswered until Besicovitch constructed such a function in 1925 [2]. A construction can be found in Jeffery [5; p. 172]. The construction is geometrical, but complicated. An arithmetical construction can be found in Morse [9]. Such functions are not typical, however, as Saks' result shows.

A quick word about the category proofs. The standard proofs are relatively straightforward. One shows directly that the class of continuous functions which possess at even one point a derivative can be written as a countable union of nowhere dense sets. For example, if E_n denotes the set of functions f in $C[0, 1]$ such that for some x in $[0, 1 - 1/n]$ the inequality $|f(x + h) - f(x)| \leq nh$ holds for all h satisfying $0 < h < 1 - x$, then it is not difficult to verify that each E_n is nowhere dense by verifying that it is closed and has a dense complement. It is also easy to see that if f has a finite right derivative at some point, then $f \in \bigcup_{n=1}^{\infty} E_n$. Thus the class of functions having finite right derivatives in a nonempty set is of the first category in $C[a, b]$. On the other hand, to show that the class of functions having *infinite* right-hand derivatives on a nonempty set is residual in $C[a, b]$ is harder to show. We know of no straightforward proof of this fact. What Saks [10] showed is that

this set is of the second category in every sphere of $C[a, b]$. Now, it is possible for a set to be of the second category in every sphere without this set being residual. But this is not possible if the set has the property of Baire. Thus, by showing that the set of functions with a finite or infinite derivative on a nondenumerable set is an analytic set in $C[a, b]$, and therefore has the property of Baire, Saks was able to complete the proof that this set is residual in $C[a, b]$.

3. The functions of Marcinkiewicz and Jarnik. So far we have discussed some of the better known pathologies possessed by typical continuous functions. We turn now to two other kinds of typical continuous functions which suggest even more pathology.

In 1935 Marcinkiewicz [6] proved that there exists a continuous function Φ with the property that to every measurable function f there corresponds a subsequence $\{h_k\}$ of the sequence $\{1/n\}$ such that

$$(1) \quad \lim_{k \rightarrow \infty} \frac{\Phi(x + h_k) - \Phi(x)}{h_k} = f(x) \text{ almost everywhere.}$$

Thus Φ is a (generalized) antiderivative for every measurable function f . As one might expect, Φ is typical — the class of all functions with this property is residual in $C[a, b]$. If we consider only the case where f is a constant, $f \equiv \lambda$, we see that the function Φ has the property that every real number λ is a derived number of Φ at almost every point; in fact, not just a derived number, but a derived number relative to a fixed sequence $\{h_k\}$. This result shows that while a typical function in $C[a, b]$ is very wildly behaved, it does possess a certain type of regularity — for each real number λ , its behavior near any point outside some null set is similar to its behavior near any other point relative to the sequence $\{h_k\}$.

By considering only constant functions, f , we have given up a good deal of Marcinkiewicz's result. But we shall gain in another direction. Let us write the expression (1) in a somewhat different form, at the same time requiring that $f \equiv \lambda$:

$$(2) \quad \lim_{\substack{y \rightarrow x \\ y \in \{x + h_k\}}} \frac{\Phi(y) - \Phi(x)}{y - x} = \lambda \text{ almost everywhere.}$$

This displays the fact that $y \rightarrow x$ while restricted to the set $\{x + h_k\}$.

Can we fatten up the set $\{x + h_k\}$ to some larger set $E(x)$ so that (2) remains valid whenever $y \rightarrow x$, $y \in E(x)$? Jarnik [4] has shown that there exist continuous functions Φ with the property that for almost every x and for every real number λ there exists a set $E(x)$ having right upper density 1 at x such that

$$\lim_{\substack{y \rightarrow x \\ y \in E(x)}} \frac{\Phi(y) - \Phi(x)}{y - x} = \lambda.$$

The statement that $E(x)$ has right upper density 1 at x means that

$$\overline{\lim}_{h \rightarrow 0} h^{-1} \mu(E(x) \cap [x, x+h]) = 1.$$

Thus $E(x)$ is “fat” near x , relative to some net of intervals $[x, x+h]$ contracting to x . We say that λ is an *essential derived number* of Φ at x .

Once again, the class of functions with this property is residual in $C[a, b]$. Thus the typical continuous function has every real number as an essential derived number at almost every point. What could be more regular than that?

We close with three remarks.

REMARK 1. It follows from Lusin’s theorem that any measurable function is the almost everywhere limit of a sequence of continuous functions. Any Marcinkiewicz type function can be used to give rise to such a sequence. Thus if f is measurable and if Φ is a Marcinkiewicz type function for which (1) holds, then $\Phi_k(x) \rightarrow f(x)$ almost everywhere, where

$$\Phi_k(x) = \frac{\Phi(x + h_k) - \Phi(x)}{h_k}.$$

Each function Φ_k is clearly continuous.

REMARK 2. Every continuous function is differentiable when restricted to a suitable subset of $[a, b]$. In fact, it was shown [3] that if f is defined and continuous on a set containing a nonempty perfect set P , there exists a nonempty perfect set $Q \subset P$ such that the restriction of f to Q is infinitely differentiable. (Here we allow the possibility that the derivatives take infinite values.)

REMARK 3. We can sum up by saying that the typical continuous function has at no point a finite or infinite derivative or a finite one-sided derivative. It has infinite one-sided derivatives on a nondenumerable set. Every real number λ is an essential derived number a.e.

The author was supported by NSF grant GP 18968.

References

1. S. Banach, Über die Baire’sche Kategorie gewisser Funktionenmengen, *Studia Math.*, 3 (1931) 174–179.
2. A. S. Besicovitch, Diskussion der stetigen Funktionen im Zusammenhang mit der Frage über ihre Differentierbarkeit, *Bull. Acad. Sci. de Russie*, 19 (1925) 527–540.
3. A. M. Bruckner, J. G. Ceder and M. L. Weiss, On the differentiability structure of real functions, *Trans. Amer. Math. Soc.* 142 (1969) 1–13.
4. V. Jarník, Sur les nombres dérivés approximatifs, *Fund. Math.*, 22 (1934) 4–16.
5. R. Jeffery, The Theorem of Functions of a Real Variable, *Math. Expositions No. 6*, University of Toronto Press, 1969.
6. J. Marcinkiewicz, Sur les nombres dérivés, *Fund. Math.*, 24 (1935) 305–308.
7. S. Mazurkiewicz, Sur les fonctions non dérivables, *Studia Math.*, 3 (1931) 92–94.

8. M. Mikolás, Construction des familles de fonctions partout continues non dérivables, Acta. Sci. Math. Szeged, 17 (1956) 49–62.

9. A. P. Morse, A continuous function with no unilateral derivatives, Trans. Amer. Math. Soc., 44 (1938) 496–507.

10. S. Saks, On the functions of Besicovitch in the space of continuous functions, Fund. Math., 19 (1932) 211–219.

REPRESENTING A FINITE BOREL MEASURE IN TERMS OF ITS DISTRIBUTION FUNCTION

J. J. HIGGINS, University of Missouri-Rolla

A real-valued function F defined on $(-\infty, \infty)$ which is bounded, non-decreasing, continuous from the right, and which satisfies the condition that $F(x) \rightarrow 0$ as $x \rightarrow -\infty$ is called a **distribution function**. Associated with a finite Borel measure μ defined on the Borel subsets of $(-\infty, \infty)$ is the distribution function $F(x) = \mu[(-\infty, x]]$. We consider the problem of giving an explicit expression for μ in terms of its distribution function F and Lebesgue measure λ . When μ has a continuous distribution function F , we see immediately that $\mu[(a, b]] = \lambda(F[(a, b]])$ for all $a < b$. Will the equation $\mu(A) = \lambda[F(A)]$ hold for all Borel sets A ? The answer is yes, provided F is continuous. Rather than proving this result directly, we prove a somewhat more general result to take into account the possibility that F has discontinuities. The result seems to give a clear picture of the way in which Borel measures are constructed.

We first establish two lemmas.

LEMMA 1. *Let $f: (-\infty, \infty) \rightarrow (-\infty, \infty)$ be a non-decreasing function. For each real number y , let $C_y = \{x: f(x) = y\}$, and let $D = \{y: C_y \text{ contains more than one point}\}$. Then D is a countable set. Furthermore, if A and B are disjoint sets, then $f(A) \cap f(B)$ is a countable set.*

Proof. As a consequence of the monotonicity of f , if C_y contains more than one point, then C_y contains an open interval. Thus each member of the collection $\{C_y\}_{y \in D}$ contains a rational number r_y ; the numbers r_y are distinct since the sets of the collection are pairwise disjoint. It follows that D is a countable set. If A and B are disjoint sets, then $f(A) \cap f(B) \subset D$; consequently, $f(A) \cap f(B)$ is a countable set.

LEMMA 2. *If F is a distribution function, then $\lambda[F(\cdot)]$ is a finite Borel measure.*

Proof. First we show that $F(A)$ is a measurable set for all Borel sets A . To this end, consider the collection $\Omega = \{A: F(A) \text{ is a measurable set}\}$. Let x_1, x_2, \dots denote the discontinuities of F . Let E_i equal either $(F(x_i^-), F(x_i))$ or $[F(x_i^-), F(x_i))$, depend-

ing on whether $F(x_i^-)$ is in the range of F or not. Then

$$F([a, b]) = [F(a), F(b)] \setminus \bigcup_{i=1}^{\infty} E_i.$$

It follows that Ω contains all intervals. Clearly Ω contains all countable unions of its members. From the equality

$$F(A^c) = [(F(A))^c \cap F((-\infty, \infty))] \cup [F(A) \cap F(A^c)]$$

and since $F(A) \cap F(A^c)$ is a countable set, it follows that Ω contains all complements of its members. Thus Ω is a sigma-algebra of subsets of $(-\infty, \infty)$ containing all intervals, and therefore contains all Borel sets.

From what has been shown, we see that $\lambda[F(\cdot)]$ is a nonnegative Borel-set function which assigns value zero to the empty set, leaving only the countable additivity to be established. Suppose $\{A_i\}_{i=1}^{\infty}$ is a sequence of pairwise disjoint Borel sets. The countable additivity of $\lambda[F(\cdot)]$ follows from Lemma 1 which gives us the fact that the sets of the sequence $\{F(A_i)\}_{i=1}^{\infty}$ are pairwise disjoint except on a set of Lebesgue measure zero. It follows that $\lambda[F(\cdot)]$ is a finite Borel measure.

THEOREM. *Let μ be a finite Borel measure whose distribution function is F . Let x_1, x_2, \dots denote the discontinuities of F , and let $j(x_i) = F(x_i) - F(x_i^-)$. Let $F_c(x) = F(x) - \sum_{x_i \leq x} j(x_i)$. Then for all Borel sets A ,*

$$(1) \quad \mu(A) = \lambda[F_c(A)] + \sum_{x_i \in A} j(x_i).$$

Proof. It is well known that F_c is a continuous distribution function (see [1] Chapter 1). The right-hand side of (1) is therefore a Borel measure, being the sum of two such measures, which agrees with μ on the intervals. The equality for all Borel sets follows from the familiar uniqueness theorem for measures (see [1], Theorem 2.2.3).

From this theorem, we see that once the properties of Lebesgue measure are established and once some of the properties of a distribution function are established, the construction of an arbitrary finite Borel measure is immediate. (For an alternate approach to the problem of constructing Borel measures, see [2], page 261.) Furthermore, we have $F_c = F_{ac} + F_s$, where F_{ac} is an absolutely continuous distribution function and F_s is a continuous, singular distribution function (see [1], chapter 1). From this we have

$$\lambda[F_c(A)] = \int_A F'_{ac}(x) dx + \lambda[F_s(A)].$$

The term $\lambda[F_s(A)] + \sum_{x_i \in A} j(x_i)$ is the singular part of the Lebesgue decomposition of $\mu(A)$, and $\int_A F'_{ac}(x) dx$ is the absolutely continuous part.

References

1. Kai Lai Chung, *A Course In Probability Theory*, Harcourt, Brace and World, New York, 1968.
2. H. L. Royden, *Real Analysis*, second edition, Macmillan, New York, 1968.

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to Shirley Hill, Department of Mathematics, University of Missouri, Kansas City, MO 64110, or to Paul Mielke, Department of Mathematics, Wabash College, Crawfordsville, IN 47933.

USING STUDENT-TUTORS IN PRECALCULUS INSTRUCTION

T. A. EISENBERG and J. B. BROWNE, Northern Michigan University

The purpose of this paper is to discuss a method of instruction which seems to be effective in handling courses with large enrollments which are primarily aimed at skill development.

The catalog description of the course we are describing is: Math 100 BASIC MATHEMATICS 4 credits

A study of the fundamental operations of algebra, factoring, graphing, linear equations, and an introduction to exponents, radicals, and quadratic equations.

The text used [1] contains most topics found in an elementary algebra course offered in high school. The course usually has an enrollment of 200 to 250 students.

We take the position that the purpose of a substantial group of courses offered by a mathematics department under the rubric of precalculus have as their main objective—skill development. The purpose of these courses is to eradicate a student's deficiencies and to enable him to perform a predetermined set of desired skills. Such precalculus courses are not ends within themselves, but are designed to prepare a student to handle whatever mathematics he will encounter in his other fields of endeavor. Survey, appreciation, and teacher education courses in mathematics which do not have at their core skill development, belong to a different group of precalculus offerings.

Because of a high failure rate in Math 100 we decided to change the classroom format of the course. Commencing in the fall semester of the '71-'72 academic year, a major switch from a large lecture only format to a lecture recitation format

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

ASSOCIATE EDITORS: JOSHUA BARLAZ, ERIC S. LANGFORD. COLLABORATING EDITORS: LEONARD CARLITZ, GULBANK D. CHAKERIAN, HASKELL COHEN, S. ASHBY FOOTE, ISRAEL N. HERSTEIN, MURRAY S. KLAMKIN, DANIEL J. KLEITMAN, ROGER C. LYNDON, MARVIN MARCUS, CHRISTOPH NEUGEBAUER, ALBERT WILANSKY, AND UNIVERSITY OF MAINE PROBLEMS GROUP: EARL M. L. BEARD, GEORGE S. CUNNINGHAM, CLAYTON W. DODGE, OSKAR FEICHTINGER, WILLIAM R. GEIGER, RAMESH GUPTA, GARY HAGGARD, PHILIP M. LOCKE, JOHN C. MAIRHUBER, CURTIS S. MORSE, GRATTAN P. MURPHY, EDWARD S. NORTHAM AND WILLIAM L. SOULE, JR.

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before September 30, 1973.

An asterisk () means neither the proposer nor the editors supplied a solution.*

E 2420. *Proposed by E. T. H. Wang, University of Waterloo, Canada*

Find all triangles with integral sides, each of which has its perimeter numerically equal to its area.

E 2421. *Proposed by R. C. Buck, University of Wisconsin*

- (i) Show that if G is a group in which the map $x \rightarrow x^3$ is a monomorphism (1:1 homomorphism, but not necessarily onto), then G is abelian.
- (ii) Exhibit a non-abelian group for which $x \rightarrow x^4$ is an automorphism.
- (iii) Are there examples for every other exponent > 4 ?

E 2422*. *Proposed by E. M. Reingold, University of Illinois, Urbana*

Two rectangles are *incomparable* if neither can be placed inside the other when they are aligned so that corresponding sides are parallel. Prove or disprove: No rectangular region can be tiled with mutually incomparable rectangles.

E 2423. *Proposed by Lyles Hoshek, Monterey Park, California, and B. M. Stewart, Michigan State University*

Let there be given a plane convex quadrilateral of area A . Divide each of its four

sides into n equal segments and join the corresponding points of division of opposite sides, forming n^2 smaller quadrilaterals. Prove: (a) the n smaller quadrilaterals in any diagonal (ordinary or broken) have a composite area equal to A/n ; (b) The composite area of any row of smaller quadrilaterals and its complementary row (row i and row $n+1-i$) is equal to $2A/n$. (In particular, if n is odd this implies that the composite area of the middle row is A/n .)

E 2424. *Proposed by P. J. Murray, Westminster College and independently by E. T. H. Wang, University of Waterloo*

Let S_n denote the set of all permutations of the first n natural numbers. We can define a metric on S_n as follows: If $\sigma, \tau \in S_n$, then $d(\sigma, \tau) = \sum_{i=1}^n |\sigma(i) - \tau(i)|$. What possible numerical values can d assume?

E 2425. *Proposed by C. A. Nicol, University of South Carolina*

For each positive real number x let $\pi_2(x)$ denote the number of twin primes not exceeding x . Show that

$$\pi_2(x) = 2 + \sum_{\gamma \leq n \leq x} \sin \frac{\pi}{2}(n+2) \left[\frac{n!}{n+2} \right] \cdot \sin \frac{\pi}{2} n \left[\frac{(n-2)!}{n} \right],$$

where brackets denote the greatest integer function.

SOLUTIONS OF ELEMENTARY PROBLEMS

Pell and the Regular n -Simplex

E 2294 [1971, 405]. *Proposed by Douglas Lind, Stanford University*

For what n does the regular n -simplex of side 1 have rational height?

Solution by Michael Goldberg, Washington, D.C. The radius r of an inscribed hypersphere of an n -dimensional simplex of unit edges is given by $r = \{1/2n(n+1)\}^{\frac{1}{n+1}}$. See H. S. M. Coxeter, *Regular Polytopes*, table on p. 295. The height is $(n+1)r$. Hence, the height is rational if r is rational, that is, if $2n(n+1) = t^2$. Since n and $(n+1)$ have no factors in common, then either

$$n = 2x^2 \quad \text{and} \quad (1) \quad 2x^2 + 1 = y^2,$$

or

$$n = u^2 \quad \text{and} \quad (2) \quad 2u^2 - 1 = v^2.$$

Equations (1) and (2) are Pell equations whose solutions can be obtained by familiar methods, e.g. Dickson, *Introduction to the Theory of Numbers*, p. 114. The least positive solution of (1) is $(x, y) = (2, 3)$. Hence the infinite sequence of solutions of (1) may be obtained by the generating equation

$$(y + \sqrt{2}x) = (3 + 2\sqrt{2})^k, \quad k = 1, 2, 3, \dots,$$

and $(x, y) = (2, 3), (12, 17), (70, 99)$, etc. Similarly, the infinite sequence of solutions (u, v) is obtained as $(u, v) = (5, 7), (29, 41), (169, 239)$, etc. Both sets of solutions are found as rational approximations of $\sqrt{2}$ which are equal to the fractions y/x and v/u . Hence $n = 8, 49, 288, 1681, 9800, 57121$, etc.

Also solved by R. T. Bumby, Cal Poly Solutions Group, R. J. Dickson, J. W. Grossman, Heiko Harborth (Germany), Ralph Jones, D. C. Kay, Harry Lass & W. S. Sinclair, D. C. B. Marsh, Henrik Meyer (Denmark), David Singmaster (England), John Stout, E. W. Trost (Switzerland), University of Toronto Geometry Students, one unidentified solver, and the proposer.

Two Conditionally Convergent Infinite Series

E 2361 [1972, 662]. *Proposed by Richard Johnsonbaugh, Morehouse College*

Prove that the following series converge conditionally:

$$\sum_{n=1}^{\infty} (-1)^n (n^{1/n} - 1) \text{ and } \sum_{n=1}^{\infty} (-1)^n [e - (1 + 1/n)^n].$$

Solution by Charles Chouteau, West Virginia State College. Convergence follows since, for $n \geq 3$, $n^{1/n}$ is a monotonic decreasing sequence converging to 1 (consider that $x^{1/x} = \exp[(1/x) \ln x]$), and $(1 + 1/n)^n$ is a sequence of positive terms monotonically increasing to e .

To show that $\sum_{n=1}^{\infty} (n^{1/n} - 1)$ diverges, we use $x - 1 \geq \ln x$ for $x > 0$, with equality if and only if $x = 1$. We have

$$\sum_{n=1}^{\infty} (n^{1/n} - 1) \geq \sum_{n=1}^{\infty} \ln n^{1/n} = \sum_{n=1}^{\infty} (1/n) \ln n,$$

which diverges by the integral test. To show that

$$\sum_{n=1}^{\infty} [e - (1 + 1/n)^n]$$

diverges, it suffices to show that $e - (1 + 1/n)^n \geq 1/(2n)$. We have

$$e = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!} + \cdots,$$

$$\left(1 + \frac{1}{n}\right)^n = 1 + 1 + \left(1 - \frac{1}{n}\right)/2! + \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)/3! + \cdots + \left(\frac{1}{n}\right)^n,$$

whence we have

$$e - (1 + 1/n)^n = (1/n)/2! + [1 - (1 - 1/n)(1 - 2/n)]/3! + \cdots.$$

Clearly all terms on the right of this last equation are positive, so their sum is greater than $1/2n$.

Also solved by M. T. Bird, D. R. Breach (New Zealand), Martin Burger, B. R. Caine, John Christopher, J. Gilles (Israel), M. G. Greening (Australia), Ellen Hertz, Michael Kostreva, Lew Kowarski,

P. A. Lindström, Z. C. Motteler, Roger Opp, G. R. Phillips, M. R. Railkar (India), W. C. Sitarick, Allen Stenger, T. J. Trent, M. K. Vamanamurthy (New Zealand), Charles Wexler, R. H. Yarbrough, P. H. Young, and the proposer.

On Spherical Triangles

E 2363 [1972, 663]. *Proposed by Hüseyin Demir, Middle East Technical University, Ankara, Turkey*

Characterize pairs of spherical triangles ABC and $A'B'C'$ for which $A' = a$, $B' = b$, $C' = c$, $A = a'$, $B = b'$, $C = c'$.

Solution by M. G. Greening, University of New South Wales, Australia. For any spherical triangle $A'B'C'$ we have:

$$(1) \quad \cos a' = \cos b' \cos c' + \sin b' \sin c' \cos A$$

and the two others following from the permutations (a, b, c) , (a, c, b) . So

$$(2) \quad \cos A = \cos B \cos C + \sin B \sin C \cos a,$$

and so on. But from consideration of the polar triangle of ABC ,

$$(3) \quad \cos A = -\cos B \cos C + \sin B \sin C \cos a.$$

Then $\cos B \cos C = \cos C \cos A = \cos A \cos B = 0$ and we have, say, $A = B = \pi/2$, yielding $a = b = \pi/2$ from (2). Also $\cos C = \cos c$, which must give $C = c$ as $c > 0$, $C < \pi$. Consequently

$$A' = B' = a' = b' = \pi/2 \text{ and } c' = C' = c = C,$$

so that the two triangles are necessarily congruent.

Also solved by Michael Goldberg, Lew Kowarski, Clellie Oursler & Eric Sturley, and the proposer.

Sylvester, Elliott, and Motzkin

E 2365 [1972, 663]. *Proposed (part I) by Erwin Just and Kenneth Fogarty, Bronx Community College, and (part II) by J. B. Wilker, University of Toronto*

I. Let S be a finite set of points in the plane in which no three points are collinear and not all points are concyclic. Define a *common* point of S to be a point which lies on some circle which passes through precisely two other points of S . Must each point of S be a common point?

II. Let S be a set of four or more points lying on a sphere but not on a circle. Prove that each point of S is on some circle containing precisely two of the other points of S .

Solution by M. G. Greening, University of New South Wales. Given n points not all of which are collinear there is always a straight line containing two but not three of the points. (This is Sylvester's problem of collinear points, dealt with in Coxeter, *Introduction to Geometry*, 1st or 2nd edition, pp. 65 and 181.)

I. Invert with respect to a circle with center $s \in S$. Then the image of $S - \{s\}$ contains two points t, u not collinear with any other image point nor collinear with s , by Sylvester's problem. The preimage of the line $t'u'$ must be a circle through s, t and u only, proving that s , and thus every point of S , is a common point.

II. By the finiteness of S there is a point $p \notin S$ such that its antipodal point $q \notin S$ and neither p nor q lies on any circle through three or more points of S . Then stereographic projection of S onto π , the tangent plane at p , maps S onto S' , which then satisfies the conditions of part I. Consequently each s' of S' lies on a circle t' through three points of S' but not through four. The bijective property of the projection and the choice of p then prove that each $s \in S$ lies on a circle t through exactly three points of S .

Also solved by A. Bruen, M. R. Railkar (India), R. R. Rottenberg (Israel), and the proposers.

Editor's Comment. In addition to the references listed on p. 65 of Coxeter, Rottenberg lists P. D. T. A. Elliott who, in his paper *On the number of circles determined by n points* (Acta Mathematica Academiae Scientiarum Hungaricae 18 (1967), 181–188), proved that in any set of $n > 3$ points in the Euclidean plane not all on a circle or line, each point lies on at least $2(n-1)/21$ circles containing exactly two other points of the set. This proves part I. Part II is proved in T. Motzkin, *The lines and planes connecting the points of a finite set*, Trans. Amer. Math. Soc. 70 (1951) 451–464. More generally, he uses “plane” and “convex surface containing no straight line” in place of our “circle” and “sphere” respectively.

Not True for $n = 8$

E 2366* [1972, 663]. *Proposed by B. P. Gill, Demarest, N.J.*

Let V be the set of vertices of a regular $2n$ -gon and let A^* and B^* be the convex n -gons whose vertices are subsets A and B of V . If the set of lengths of all chords with both ends in A (with each chord length being counted according to its multiplicity) is identical to the like set for B , then is B^* necessarily congruent to either A^* or $(V \setminus A)^*$?

Solution by Herbert Taylor, Skymount, Va. Not necessarily, as shown by the easily verified counterexample for $n = 8$ where A consists of vertices 1, 2, 3, 5, 6, 9, 11, and 13, and B consists of vertices 1, 2, 3, 5, 9, 11, 13, and 14 of a 16-gon. In each figure one finds 3, 4, 3, 5, 3, 4, 3, 3 chords of the seven possible lengths, listed in order of increasing length.

Editorial Note. Two attempted solutions were also received, one utilizing modular congruences to obtain an algebraic criterion to determine for which n the problem is or is not true. A correct algebraic solution is solicited.

Square Numbers in a Recursive Sequence

E 2367 [1972, 772]. *Proposed by Erwin Just, Bronx Community College*

Let F_n be the n th term of the sequence defined by

$$F_n = -F_{n-1} - 2F_{n-2}, \quad F_1 = 1, \quad F_2 = -1.$$

Prove that $2^{n+1} - 7F_{n-1}^2$ is a perfect square.

I. *Solution by Trygve Breiteig, Kristiansand Laerarskole, Norway.* By induction on n we prove that

$$2^{n+1} - 7F_{n-1}^2 = (2F_n + F_{n-1})^2.$$

Clearly this equation holds for $n = 2$, so we assume it correct for n , $n \geq 2$. Then we have

$$\begin{aligned} (2F_{n+1} + F_n)^2 &= (-F_n - 4F_{n-1})^2 = F_n^2 + 8F_nF_{n-1} + 16F_{n-1}^2 \\ &= 2(4F_n^2 + 4F_nF_{n-1} + F_{n-1}^2) + 14F_{n-1}^2 - 7F_n^2 \\ &= 2(2^{n+1} - 7F_{n-1}^2) + 14F_{n-1}^2 - 7F_n^2 = 2^{n+2} - 7F_n^2. \end{aligned}$$

II. *Solution by C. R. Wall, University of South Carolina.* By the recurrence relation we have $F_0 = 0$ and

$$\begin{bmatrix} F_n & F_{n+1} \\ F_{n-1} & F_n \end{bmatrix} \cdot \begin{bmatrix} 0 & -2 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} F_{n+1} & F_{n+2} \\ F_n & F_{n+1} \end{bmatrix}$$

so that

$$\begin{bmatrix} F_n & F_{n+1} \\ F_{n-1} & F_n \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 & -2 \\ 1 & -1 \end{bmatrix}^{n-1}.$$

We take determinants to obtain $F_n^2 - F_{n+1}F_{n-1} = 2^{n-1}$, multiply both sides by 4, substitute $-F_{n+1} = F_n + 2F_{n-1}$, and subtract $7F_{n-1}^2$ from both sides to obtain

$$(2F_n + F_{n-1})^2 = 2^{n+1} - 7F_{n-1}^2.$$

Also solved by the proposer and sixty-four others, many of whom submitted more than one solution. Most of the solutions made use of formulas for recurrence functions or difference equations. There were a number of generalizations, the trend of which can be inferred from the following references (given by our readers):

B. W. Jones, *The Theory of Numbers*, p. 73, ex. 5.

G. Polya, *Mathematical Discovery*, pp. 108-109, 196-197.

J. Arkin, *Convergence of the coefficients in a recurring power series*, Fibonacci Quarterly, vol. 7, no. 1, February 1969.

M. Ward, *The arithmetical theory of linear recurring series*, Trans. Amer. Math. Soc., 35 (1933), 600-628.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers — The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before September 30, 1973. Contributors (in the United States) who desire acknowledgement of receipt of their solutions are asked to enclose self-addressed, stamped postcards.

5916. *Proposed by J. G. Mauldon, Amherst College*

Let p and q be coprime integers with $0 < p < q$ and let $f(\cdot)$ be a non-negative function defined on $\{0, 1, 2, \dots, p\}$ such that $f(0) = 0$, $f(p) = 1$ and, whenever $\sum_{i=1}^k m_i = q$ with $0 \leq m_i \leq p$ ($i = 1, 2, \dots, k$), we have $f(m_k) \leq \sum_{i=1}^{k-1} f(m_i)$. Is it necessarily true that $\max \{f(n): n = 0, 1, 2, \dots, p\} < 3q$?

5917. *Proposed by Andrzej Ehrenfeucht, University of Colorado*

Let $f: [0, 1] \rightarrow \mathbb{R}^2$ be a curve such that $f(0) = (0, 0)$, $f(1) = (1, 0)$ and for every $0 \leq r < s < t \leq 1$ we have

$$\max \{ \|f(r) - f(s)\|, \|f(s) - f(t)\| \} \leq \|f(r) - f(t)\|,$$

i.e., f never approaches any point through which it went and does not move away from any point through which it will pass. Prove that such curves have length and that those lengths are bounded. Is $2\pi/3$ an upper bound?

5918. *Proposed by S. W. Golomb, University of Southern California*

Prove or disprove: Let C be a collection of distinct subsets of the positive integers which are totally ordered by the inclusion relation. Clearly C must be either a finite or a countable collection.

5919. *Proposed by L. A. Harris, University of Kentucky*

Show that if x and y are two vectors in a complex Hilbert space and if n is a positive integer, then

$$\sum_{k=1}^{2n} \|x + \lambda^k y\|^{2n} \leq n \binom{2n}{n} (\|x\|^{2n} + \|y\|^{2n}),$$

where $\lambda = \exp(i\pi/n)$. Note that this generalizes the parallelogram law.

5920. *Proposed by L. C. Washington, Princeton University*

Do there exist nonzero compact (Hausdorff) topological vector spaces over infinite fields?

5921. *Proposed by Paul Cohn, Bedford College, London, England*

Over any commutative field (e.g., the complex numbers), find two square matrices

of the same order, A and B , such that every matrix in the pencil $\lambda A + \mu B$ is nilpotent, but A and B cannot be simultaneously triangularized.

SOLUTIONS OF ADVANCED PROBLEMS

σ -algebras in $X \times X$

5845 [1972, 307]. *Proposed by J. A. Johnson, Oklahoma State University*

Let X be an uncountable set and \mathcal{A} the smallest σ -algebra of subsets of $X \times X$ containing all sets of the form $A \times B$ where $A \subset X$, $B \subset X$. Does \mathcal{A} contain all subsets of $X \times X$?

Solution by Manfred Stoll and R. L. Taylor, University of South Carolina. If the cardinality of X is greater than the cardinality of the continuum, an example of a subset of $X \times X$ which is not in \mathcal{A} is given on p. 261 of Halmos, *Measure Theory*, Van Nostrand Co., Inc. 1950. If one assumes the continuum hypothesis, the question has been answered affirmatively for all sets X with cardinality less than or equal to the continuum by B. V. Rao, *On discrete Borel spaces and projective sets*, Bull. A.M.S., 75 (1969) 614.

Also solved by E. J. Braude, R. E. Frankfurt, S. J. Garland, A. A. Jagers (Netherlands), Douglas Lind, Michael McCoy, and Mary Powderly.

Generalization of the Riemann-Lebesgue Lemma

5846 [1972, 307]. *Proposed by H. Kestelman, University College, London, England*

If $f \in L(0, \infty)$ and $I(\lambda)$ is, for each positive λ , a subinterval of $(0, \infty)$, then $\lim_{\lambda \rightarrow \infty} \int_{I(\lambda)} f(t) \cos \lambda t dt = 0$. If $I(\lambda)$ is assumed only to be the union of a finite set of intervals, the result is false.

Solution by David Borwein, University of Western Ontario. (i) Suppose $I(\lambda) = (a_\lambda, b_\lambda)$, and let

$$S_\lambda = \int_{I(\lambda)} f(t) \cos \lambda t dt.$$

Then

$$S_\lambda = \int_{a_\lambda}^{b_\lambda} f(t) \cos \lambda t dt = - \int_{a_\lambda - \pi/\lambda}^{b_\lambda - \pi/\lambda} f(t + \pi/\lambda) \cos \lambda t dt;$$

and so

$$\begin{aligned}
2|S_\lambda| &\leq \int_{a_\lambda}^{b_\lambda} |f(t) - f(t + \pi/\lambda)| dt + \int_{a_\lambda}^{a_\lambda + \pi/\lambda} |f(t)| dt \\
&\quad + \int_{b_\lambda}^{b_\lambda + \pi/\lambda} |f(t)| dt \\
&= o(1) \text{ as } \lambda \rightarrow \infty.
\end{aligned}$$

(ii) Suppose now that

$$I(n) = \bigcup_{r=1}^n \left(\frac{(4r-1)\pi}{2n}, \frac{2r\pi}{n} \right), \quad n = 1, 2, \dots,$$

and that f is the characteristic function of $(0, 2\pi)$. Then

$$\int_{I(n)} f(t) \cos nt dt = \sum_{r=1}^n \frac{1}{n} = 1.$$

Also solved by L. E. Clarke (England), A. A. Jagers (Netherlands), O.P. Lossers (Netherlands), J. E. Mazo & G. J. Foschini, and the proposer.

Unique Fixed Points

5847 [1972, 307]. *Proposed by Joe Beasley, Prairie View A&M College*

X is a complete metric space and $T: X \rightarrow X$ is a function with the following conditions:

- (1) There is a sequence $\{x_n\} \in X$ such that $d(x_n, T(x_n)) \rightarrow 0$.
- (2) $t: X \rightarrow R$ defined by $t(x) = d(x, T(x))$ is lower semicontinuous.
- (3) $d(T(x), T(y)) \leq ad(x, T(x)) + bd(y, T(y)) + cd(x, y)$, where a, b, c are positive numbers and $c < 1$.

Show that (A) T has a unique fixed point, and (B) none of conditions (1), (2) or (3) can be omitted.

Note. This has already appeared in this Department. See 5775 [1972, 95].

New solutions were submitted by C. Crofts & W. R. Woodward, W. S. Hall, Rik Harris & Frank Oliva, Joel Levy, O. P. Lossers (Netherlands), Mark Rowles, T. Šalát (Czechoslovakia), V. M. Sehgal, K. D. Stroyan, R. K. Tamaki, H. C. Wentz, and the proposer.

Zeros of $z^n + z^m - 1$

5848 [1972, 399]. *Proposed by A. Smith, Carleton University, Ottawa*

Let m and n be positive coprime integers. Find the number of zeros of the function $z^n + z^m - 1$ which lie inside the unit circle.

Solution by Nancy M. Huddleston and C. C. Rousseau, Memphis State University. Let $f(z) = z^n + z^m - 1$ and let N denote the number of zeros of $f(z)$ which lie within the unit circle. If $z = e^{i\theta}$, then

$$(1) \quad \operatorname{Re} f(e^{i\theta}) = 2 \cos \left(\frac{(n+m)\theta}{2} \right) \cos \left(\frac{(n-m)\theta}{2} \right) - 1$$

and

$$(2) \quad \operatorname{Im} f(e^{i\theta}) = 2 \sin \left(\frac{(n+m)\theta}{2} \right) \cos \left(\frac{(n-m)\theta}{2} \right).$$

From (1) and (2) it follows that $f(z)$ has zeros on the unit circle only if $n + m \equiv 0 \pmod{6}$. If $n + m \not\equiv 0 \pmod{6}$, the principle of the argument can be applied and N is equal to the number of times $w = f(e^{i\theta})$ encircles the origin as θ goes from 0 to 2π . When $\theta = 0$, $\operatorname{Re} w = 1$ and $\operatorname{Im} w = 0$. Setting up the conditions

$$(3) \quad \operatorname{Re} f(e^{i\theta}) > 0, \quad \operatorname{Im} f(e^{i\theta}) = 0,$$

and using (1) and (2), we find that $d/d\theta [\operatorname{Im} f(e^{i\theta})] > 0$ whenever (3) is satisfied. It follows that N is equal to the number of values of θ ($0 < \theta \leq 2\pi$) such that (3) is satisfied. Since, in order to satisfy (3), it is necessary to take $(n+m)\theta/2 = k\pi$, N is equal to the number of values of k ($0 < k \leq n+m$) such that

$$(4) \quad (-1)^k \cos \left(\frac{(n-m)k\pi}{n+m} \right) > \frac{1}{2}.$$

Equivalently, we want the number of integers k ($0 < k \leq n+m$) such that

$$(5) \quad \left| \frac{(n-m)k\pi}{n+m} - l\pi \right| < \frac{\pi}{3}$$

with k and l both odd or both even. Letting $k+l=2r$ and $k-l=2s$, we write (5) in the form

$$(6) \quad |ns - mr| \leq \left\lceil \frac{n+m}{6} \right\rceil.$$

Since we require that $(n, m) = 1$, it follows that there exist solutions of $ns - mr = p$ for every integer p . Moreover, for each such p , there is a unique solution which satisfies the condition $0 < k \leq n+m$. Hence, if $n+m \not\equiv 0 \pmod{6}$, the number of zeros lying within the unit circle is given by

$$(7) \quad N = 2 \left\lceil \frac{n+m}{6} \right\rceil + 1.$$

If $n+m \equiv 0 \pmod{6}$, it follows easily that $f(z)$ has precisely two zeros which lie on the unit circle and these zeros are at $e^{in/3}$ and $e^{-in/3}$. For this case, one can apply

the principle of the argument to an appropriately indented contour and obtain

$$(8) \quad N = \frac{n+m}{3} - 1.$$

Also solved by Robert Breusch, O. P. Lossers (Netherlands), St. Olaf College Students, and the proposer.

Note. Breusch and Lossers express N as $2[(n+m-1)/6] + 1$ for all cases of $n+m$. The proposer notes that N is the nearest odd integer to $(n+m)/3$, and is the lesser one if $n+m \equiv 0 \pmod{6}$.

An Unsolved Approximation Question for π

5849 [1972, 399]. *Proposed by H. D. Ruderman, Hunter College High School, New York City*

For positive integers n , what is the Greatest Lower Bound for $n|\sin n|$?

Comments by F. G. Schmitt, Jr., Berkeley, Cal. If the simple continued fraction representation of π is denoted by $\pi = [a_0, a_1, a_2, \dots]$, with convergents $p_n/q_n = [a_0, a_1, \dots, a_n]$, then [1, p. 163] $|p_n/q_n - \pi| < 1/a_{n+1}q_n^2$, so

$$\begin{aligned} p_n |\sin p_n| &= p_n |\sin(p_n - q_n \pi)| = p_n \sin |p_n - q_n \pi| \\ &< p_n |p_n - q_n \pi| < p_n / a_{n+1} q_n. \end{aligned}$$

But the convergents are bounded, so if the partial quotients a_n are unbounded, then $\text{glb } n|\sin n| = 0$.

However, it is unknown if the a_n are unbounded, although Lehmer (2) comments on their "seemingly lawless" behavior, and it is known (1, Thm. 196, p. 166) that almost all real numbers have unbounded partial quotients. The latest computation of the first 21230 values of a_n (3) has failed to disclose one larger than $a_{431} = 20776$, discovered by Lehman (4). Hence, since $p_n/q_n < \pi$ for n even, the best result to date from this point of view is $\text{glb } n|\sin n| < \pi/20776$.

References

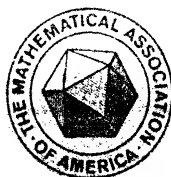
1. Hardy, G. H. and E. M. Wright, *An Introduction to Theory of Numbers*, 4th ed., Oxford, London, 1965.
2. Lehmer, D. H., *Note on an absolute constant of Khintchine*, this MONTHLY, 46 (1939) 148-152.
3. Choong, K. Y., D. E. Daykin and C. R. Rathbone, *Rational Approximations to π* , Math. Comp., 25 (1971) 387-392.
4. Lehman, R. S., *A Study of Regular Continued Fractions*, BRL Report 1066, Aberdeen Proving Ground, Maryland, February 1959.

Additional comments were submitted by Richard Gisselquist, A. A. Jagers (Netherlands), and Steven Russ.

THE AMERICAN MATHEMATICAL MONTHLY

THE OFFICIAL JOURNAL OF
THE MATHEMATICAL ASSOCIATION OF AMERICA, INC.

VOLUME 80



NUMBER 6

CODEN: AMMYAE

PART II

Papers in the Foundations of Mathematics

Number 13

of the

HERBERT ELLSWORTH SLAUGHT
MEMORIAL PAPERS

JUNE-JULY

1973

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss.

Backlog: Main Articles 12 months, Math. Notes 13 months, Research Problems 7 months, Classroom Notes 11 months, Math. Education 10 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to ALEX ROSENBERG, Department of Mathematics, Cornell University, Ithaca, N.Y. 14850; NOTES, etc.; to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D.C. 20036.

HARLEY FLANDERS, *Editor*

ALEX ROSENBERG, *Editor-Elect*

ASSOCIATE EDITORS

JOSHUA BARLAZ
E. R. BERLEKAMP
JANE W. DI PAOLA
ROBERT GILMER
RICHARD GUY
RAOUL HAILPERN

J. G. HARVEY
ERIC S. LANGFORD
P. D. LAX
ARTHUR MATTUCK
M. W. POWNALL
GIAN-CARLO ROTA

SEYMOUR SCHUSTER
J. ARTHUR SEEBACH, JR.
E. P. STARKE
LYNN A. STEEN
JAMES WENDEL

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are \$12.50. For nonmembers the subscription price is \$18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June-July, August-September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices.

Copyright © The Mathematical Association of America (Incorporated), 1973

PRINTED IN THE UNITED STATES OF AMERICA

PAPERS IN THE FOUNDATIONS OF MATHEMATICS

CONTENTS

Preface J. C. ABBOTT 1

Some Aspects of the Theory of Models ROBERT L. VAUGHT 3

What is Nonstandard Analysis? W. A. J. LUXEMBURG 38

Recursive Functions and Hierarchies HILARY PUTNAM 68

Function Theory on Some Nonarchimedean Fields . ABRAHAM ROBINSON 87

The Thirteenth

HERBERT ELLSWORTH SLAUGHT
MEMORIAL PAPER

Published as a supplement to the AMERICAN MATHEMATICAL MONTHLY

Volume 80 June-July 1973 Number 6

PREFACE

The papers in this volume of the Slaughter Papers were first presented to a Symposium at the United States Naval Academy in May of 1968. The purpose of the symposium was to provide a service to the Mathematical Community by presenting various leaders in the broad area of "Foundations of Mathematics" affording them an opportunity to explain the then current situation of their particular specialty and to provide a forum in which they could present their thoughts concerning future trends. Mathematics today and Foundations in particular is expanding so rapidly that it is difficult for even the expert to keep abreast in all but a narrow corner of a vast realm. At the same time there is a greater need than ever for the general mathematician and particularly the teacher of mathematics to keep posted on what is going on in the various fields of mathematics. It was for this purpose that four of the country's leading authorities within the general area of Foundations were invited to give their considered views within their particular specialty. These speakers were selected not only because they possess the most intimate knowledge of their field, but because they are also known as great teachers themselves, who have a desire to share their expertise with others.

The first paper on Model Theory is by Professor Robert Vaught of the University of California at Berkeley. Professor Vaught obtained his Ph. D. degree in 1954 under the direction of Alfred Tarski at the University of California at Berkeley. After four years at the University of Washington he returned to Berkeley where he has been a Professor since 1963. He served as a Fulbright Research Scholar in Amsterdam (1956-57), a senior post-doctoral NSF Fellow at UCLA (1963-64), and a Guggenheim Fellow at Zürich (1967). Aside from the theory of models Professor Vaught has made important contributions to the foundations of set theory and to the theory of decision methods.

The second paper is on Non-Standard Analysis by Professor W. A. J. Luxemburg of California Institute of Technology. Professor Luxemburg received his doctorate from Delft in 1955. He was a Fellow at Queens in 1955-56 and an Assistant Professor at Toronto (1956-58) and has been a Professor at California Institute of Technology since then. His early interests were in functional analysis with emphasis on measure and integration theory, Banach spaces, locally convex spaces and Riesz spaces as well as ordinary differential equations and topological linear spaces. More recently he has done much work in non-standard Analysis, Boolean algebra and axiomatic set theory.

The third paper in this series is on Recursive Functions by Professor Hilary Putnam of Harvard University's Department of Philosophy. Professor Putnam received his Ph. D. from the University of California at Los Angeles in 1951. He was Instructor at Northwestern University (1952-53), Assistant Professor of Philosophy at Princeton (1953-60), where he also gave seminars in logic and the philosophy of science at Swarthmore, Associate Professor, Princeton (1960-61) and

Professor of Philosophy of Science at MIT (1961–65). He has been Professor of Philosophy at Harvard since then. Professor Putnam was also a Rockefeller Foundation Research Fellow, 1951–52, Visiting Research Professor at the Minnesota Center for the Philosophy of Science (1957) and a Guggenheim Foundation Fellow, 1960–61. Professor Putnam has done research in the philosophy of science leading to articles on physical geometry, quantum mechanics, the analytic-synthetic distinction and its role in physical theory and various aspects of artificial intelligence, philosophical behaviorism and contemporary linguistic theory. His papers in logic include papers on three valued logic, decidability, set theory, recursive functions and constructable sets amongst many others.

The final paper in this volume by Professor Abraham Robinson on Recent developments in Non-Standard Analysis was presented as a supplement to Professor Luxemburg's paper at the Naval Academy in the Spring of 1970. Professor Robinson received his doctorate in 1949 from the University of London. His original training was in mechanical and aerodynamic engineering, having served as an engineer with the Royal Aircraft Establishment from 1942–1951. He was at the University of Toronto for the period 1951–57 and was Professor at Hebrew University, Israel, 1957–62. For the period 1962–67 he was Professor at the University of California at Los Angeles in Mathematics and Philosophy and went to Yale in 1967 where he has remained since. At various times he has been in a visiting capacity at Princeton, Oxford, Paris, Rome, Tübingen and Hebrew University. His early work was in wing theory and aerodynamics. Later he turned to model theory, the mathematics of algebra, mathematical logic and non-standard analysis with applications to analysis and algebra. He is at present President of the Association for Symbolic Logic.

We wish to thank the Office of Naval Research for their financial support of the Symposium.

J. C. ABBOTT

SOME ASPECTS OF THE THEORY OF MODELS

ROBERT L. VAUGHT, University of California at Berkeley

In modern mathematics many of the questions studied are of the form: does a certain mathematical structure possess a given property? In the theory of models one studies the collection of all properties or, rather, some restricted but large collection of properties. Usually these are just the properties expressible in a certain language.

For example, each of the properties of being a group, an Abelian group, or a torsion-free Abelian group is expressible in the so-called elementary language (or first-order predicate calculus). Thus, instead of saying the group \mathcal{G} is Abelian, we can say it is a model of the elementary sentence $\forall x \forall y (x \circ y = y \circ x)$. Such properties are also called elementary.

It turns out that there are results applying to *all* elementary properties which are so strong that we may have been unaware of their instances in simple cases like those above. Apparently the result holds just because of the special way the property can be expressed, and we had not previously been paying attention to that aspect of the property. Perhaps the most fundamental such result is Gödel's Compactness Theorem: *If $\sigma_1, \dots, \sigma_n, \dots$ are elementary sentences and for each n , $\{\sigma_1, \dots, \sigma_n\}$ has a model, then $\{\sigma_1, \dots, \sigma_n, \dots\}$ has a model.* (This is proved, in a more general form, in §2.)

In this article we introduce the reader with little or no background in logic to the basic results of model theory and to a selected group of advanced or recent topics as well. Specifically, in the first three sections the basic notions of model theory and the Compactness and Löwenheim-Skolem theorems, as well as their applications, are treated. In the last four sections, we discuss saturated models, ultraproducts, the ω -completeness theorem and Löwenheim-Skolem theorems for two cardinals. Of course very many important topics are omitted, but those we do cover hang together closely and give a good idea of the flavour of the subject. Full proofs are given for all theorems, except for a few results mentioned on the side.

1. The elementary language. Groups are structures of the form (G, \circ) . Some other structures of different forms are ordered sets $(A, <)$, ordered rings with unit $(A, +, \cdot, <, 1)$, etc. Some structures, such as groups with operators, even have infinitely many distinguished operations (or distinguished relations or distinguished elements). In general, then, to determine a 'similarity type' of structures, we are given an arbitrary set S of 'symbols', each of which is classified as an n -ary operation symbol (for some $n \geq 1$), or as an n -ary relation symbol, or as an individual constant. \mathfrak{A} is a structure of **similarity type** S (or **S -structure**) if $\mathfrak{A} = (A, F)$, where A is a non-empty set and F is a function on S such that for each $X \in S$, $F(X)$ is an n -ary operation over A or an n -ary relation over A (i.e., a set of n -tuples of

elements of A) or a member of A , according to which kind of symbol X is. We agree to write $X^{\mathfrak{U}}$ for $F(X)$ and to write $\mathfrak{U} = (A, X^{\mathfrak{U}})_{X \in S}$. (If S is finite and we fix a listing X_1, \dots, X_n of its members, then we can write $\mathfrak{U} = (A, X_1^{\mathfrak{U}}, \dots, X_n^{\mathfrak{U}})$ as in the examples above.) A is called the **universe** of \mathfrak{U} . We always understand that ' A ' denotes the universe of \mathfrak{U} , ' B ' that of \mathfrak{B} , etc. \mathfrak{U} determines S , so we can write $S = S(\mathfrak{U})$. Incidentally, ' c ', ' d ' will always denote individual constants, ' \mathcal{O} ' an operation symbol, and ' P ', ' Q ' relation symbols.

The **elementary language** $L(S)$ has the non-logical symbols X (for $X \in S$) and the logical symbols $\sim, \wedge, \exists, \approx, ,$ and $v_0, v_1, \dots, v_n, \dots$ (called variables), as well as $(,)$. ' u ', ' v ', ' w ' always denote variables. The symbols are put together in the obvious ways, but it will be helpful to have the precise definitions: The set of **terms** of $L(S)$ is the smallest set containing all variables and individual constants and containing $\mathcal{O}\tau_1 \dots \tau_n$ whenever it contains τ_1, \dots, τ_n and \mathcal{O} is an n -ary operation symbol in S . ' τ ' always denotes a term. If P is an n -ary relation symbol in S and τ_1, \dots, τ_n are terms then $P\tau_1 \dots \tau_n$ is an **atomic formula** of $L(S)$. (\approx is considered to be a binary relation symbol.) The set of **formulas** of $L(S)$ is the smallest set containing all atomic formulas of $L(S)$ and containing $\sim\phi$, $(\phi \wedge \psi)$, and $\exists v_k \phi$ whenever it contains ϕ and ψ . θ, ϕ, ψ always denote formulas. The precise definition points up the inductive character of being a formula, which is of basic importance in proving results about the properties corresponding to formulas.

We write $\phi \vee \psi$ to mean $\sim(\sim\phi \wedge \sim\psi)$; and $\rightarrow, \leftrightarrow, \forall v_n$, and $\exists! v_n$ (read 'there exists exactly one v_n such that') are understood similarly. But it should be remembered that in $L(S)$ itself only \sim, \wedge, \approx , and \exists occur.

For some examples, let us now consider groups as structures of the form $\mathcal{G} = (G, \circ, e)$, of type $S_0 = \{\circ, e\}$. (It is not necessary to have e , but it simplifies various formulas below.) The formula $\phi: \forall v_1(v_0 \circ v_1 \approx v_1 \circ v_0)$ has v_0 as its only **free variable**. Formulas with no free variables are called **sentences**. σ will always range over sentences.

We cannot ask whether the formula ϕ above is true (or holds) in the group \mathcal{G} until we also have at hand a particular element b of G (to be 'denoted' by v_0). Given both \mathcal{G} and b we say that ϕ is **satisfied** in \mathcal{G} when b is assigned to v_0 , and write $\mathcal{G} \vdash \phi [\overset{v_0}{b}]$, just if, in fact, $b \circ x = x \circ b$ for all $x \in G$. (The term 'elementary' or 'first-order' refers to the fact that in interpreting satisfaction in, say, \mathcal{G} , all quantified variables range over the elements of G , and not over subsets or subrelations of G (second order) or natural numbers or whatever.) In a similar way we understand the notation

$$\mathfrak{U} \vdash \psi \left[\begin{matrix} u_0 \dots u_{n-1} \\ a_0 \dots a_{n-1} \end{matrix} \right],$$

where $a_i \in A$ and u_0, \dots, u_{n-1} are distinct and include all free variables of ψ . For a sentence σ we write simply $\mathfrak{U} \vdash \sigma$ and say σ is **true** in \mathfrak{U} or \mathfrak{U} is a **model** of σ .

Strictly speaking, the general notion of satisfaction (like that of formula) is defined by an inductive procedure, with one clause for atomic formulas, and one each for \sim , \wedge , and \exists . The clause for \sim is

$$\mathfrak{A} \vdash (\sim \psi) \begin{bmatrix} u_0 \cdots u_{n-1} \\ a_0 \cdots a_{n-1} \end{bmatrix}$$

if and only if it is not the case that

$$\mathfrak{A} \vdash \psi \begin{bmatrix} u_0 \cdots u_{n-1} \\ a_0 \cdots a_{n-1} \end{bmatrix}.$$

From this example the reader should be able to supply the other clauses himself (and should do so). Actually, when operation symbols are present we must first define $\tau^{\mathfrak{A}} \begin{bmatrix} u_0 \cdots u_{n-1} \\ a_0 \cdots a_{n-1} \end{bmatrix}$, the value of the term τ when a_i is assigned to u_i ($i < n$).

This is the same as the usual 'value' of a polynomial in algebra.

When u_0, \dots, u_{n-1} are exactly v_0, \dots, v_{n-1} we write simply $\mathfrak{A} \vdash \phi[a_0, \dots, a_{n-1}]$ instead of $\mathfrak{A} \vdash \phi \begin{bmatrix} v_0 \cdots v_{n-1} \\ a_0 \cdots a_{n-1} \end{bmatrix}$. A formula ψ whose free variables are among v_0, \dots, v_{n-1} is called an *n-formula*. It creates an *n*-ary relation $\psi^{\mathfrak{A}}$ in a structure \mathfrak{A} , namely, $\psi^{\mathfrak{A}} = \{(a_0, \dots, a_{n-1}) : \mathfrak{A} \vdash \psi[a_0, \dots, a_{n-1}]\}$. An *n*-ary relation over A is called **definable** in \mathfrak{A} just if it is of this form for some ψ . Returning to the 1-formula $\phi: \forall v_1(v_1 \circ v_0 \approx v_0 \circ v_1)$, we see that for a group \mathcal{G} , $\phi^{\mathcal{G}}$ is just the center of the group (the set of elements which commute with all elements).

It is clear that we can write down a familiar sentence (the conjunction of three or four sentences) such that a structure $\mathcal{G} = (G, \circ, e)$ is a group if and only if it is a model of σ . To indicate this state of affairs we say that the property of being a group is **elementary**, or that the class of all groups is \mathcal{EC} . Consider now the property of being a torsion-free group, i.e., a group (G, \circ, e) in which

(1) for any $x \neq e$, and any positive integer k , $x^k \neq e$.

It does not appear that we can state (1) equivalently as an elementary sentence, since it has a quantifier ranging over positive integers (and indeed it can be proved impossible—see end, §2). However, we can paraphrase (1) by an infinite set Σ of sentences. Σ consists of $\forall u(u \neq e \rightarrow u \circ u \neq e)$, $\forall u(u \neq e \rightarrow u \circ (u \circ u) \neq e)$, and so on. Thus, the class \mathcal{K} of all torsion-free groups is the class $\text{Mod } \Sigma'$ of all models of a certain set $\Sigma' = \Sigma \cup \{\sigma\}$ of elementary sentences. (Being a model of a set of sentences means being a model of each.) We say \mathcal{K} is *elementary in the wider sense* or \mathcal{EC}_Δ . It turns out that such \mathcal{K} are almost as well-behaved as the elementary.

Another property of groups is that of being a torsion (or periodic) group, i.e., one in which every element is of finite order (for any a , $a^n = e$ for some positive

integer n). This property is known (see end, §2) not even to be elementary in the wider sense. However, it can be expressed by a single sentence of another language called $L_{\omega_1\omega}$, which we shall consider later. Such properties will be called $\mathcal{EC}(L_{\omega_1\omega})$.

Before giving more examples, we need to introduce another classification of properties, called \mathcal{PC}_Δ . If \mathfrak{U} is an S -structure and $S' \subseteq S$ then the **reduct** $\mathfrak{U} \upharpoonright S'$ is simply the structure $\langle A, X^{\mathfrak{U}} \rangle_{X \in S'}$. (An example is the additive group of a ring.) In the reverse direction, the structure $\mathfrak{B} = (\mathfrak{U}, Y_i)_{i \in I}$, where each Y_i is a finitary relation or operation over A or an element of A , is just obtained by adding or adjoining the Y_i to all the $X^{\mathfrak{U}} (X \in S)$; thus $\mathfrak{B} \upharpoonright S = \mathfrak{U}$. (Strictly speaking we must first make up new symbols of the appropriate type and add them to S to get an $S'' \supseteq S$ such that \mathfrak{B} is an S'' -structure.) \mathfrak{B} is called an **expansion** of \mathfrak{U} .

If \mathcal{K} is an \mathcal{EC}_Δ -class of S' structures and $S' \supseteq S$, then the class $\mathcal{K} \upharpoonright S = \{\mathfrak{U} \upharpoonright S : \mathfrak{U} \in \mathcal{K}\}$ is called a \mathcal{PC}_Δ -class (pseudo-elementary class). Clearly in defining \mathcal{K} we use an existential second-order quantifier and even more ($\mathfrak{U} \in \mathcal{K}$ if there exist relations or operations $X (X \in S' - S)$ over A such that ...). Therefore, as one would expect, it can be shown that \mathcal{PC}_Δ is a broader classification than \mathcal{EC}_Δ (see below). Nevertheless it turns out that \mathcal{PC}_Δ -properties have many of the features of elementary properties. (The notions \mathcal{EC} , \mathcal{EC}_Δ , and \mathcal{PC}_Δ were all introduced by Tarski. See [14] for an excellent general discussion.)

For some more examples, consider the class \mathcal{K} of all well-orderings $(A, <)$, and the class \mathcal{K}' of all orderings which are not well-ordered. It is trivial to verify that \mathcal{K}' is \mathcal{PC}_Δ . \mathcal{K}' is not \mathcal{EC}_Δ (see end, §2). On the other hand, to say \mathfrak{U} is well-ordered (every non-empty subset has a first element) requires universally quantifying over all subsets of A . This appears to be very strong, but of course we must beware of the possibility that some simpler, equivalent definition of \mathcal{K} can be given establishing that \mathcal{K} is, say, \mathcal{EC}_Δ or $EC(L_{\omega_1\omega})$. Proving this is not so is sometimes difficult and there are many interesting problems of this form. As regards \mathcal{K} , however, it is not \mathcal{PC}_Δ (see end of §2) and it is not (unlike 'torsion group') even $\mathcal{PC}_\Delta(L_{\omega_1\omega})$ (see [8]). There is, however, a very strong language, $L_{\omega_1\omega_1}$, which has been studied, in which being well-ordered is expressible.

We need a little more syntactical notation. If ϕ is a formula, u_0, \dots, u_{n-1} are distinct variables and $\tau_0, \dots, \tau_{n-1}$ are terms then $\phi \left(\begin{smallmatrix} u_0 \cdots u_{n-1} \\ \tau_0 \cdots \tau_{n-1} \end{smallmatrix} \right)$ is the formula obtained from ϕ by simultaneously substituting τ_i for each free occurrence of u_i ($i < n$) (after first changing bound variables so that no variables in the τ_i occur bound in ϕ). $\phi(\tau_0 \cdots \tau_{n-1})$ means $\phi \left(\begin{smallmatrix} v_0 \cdots v_{n-1} \\ \tau_0 \cdots \tau_{n-1} \end{smallmatrix} \right)$. We write $\Sigma \vdash \phi$, and call ϕ a **consequence** of Σ , if ϕ is true in all models of Σ ; if ϕ is not a sentence it is understood to be replaced by its **universalization** $(\forall u_0 \cdots u_{n-1})\phi$, where u_0, \dots, u_{n-1} are all the free variables of ϕ . ϕ is **logically valid** if $\vdash \phi$ (i.e., $0 \vdash \phi$). ' Σ ' always denotes a set of sentences.

A binary operation (for example) is clearly essentially the same thing as a ternary relation for which the sentence $(\forall u, v)\exists! w \text{ } Puvw$ holds. Moreover, it is easy to check

that any elementary sentence involving a binary operation symbol \mathcal{O} can be expressed as an elementary sentence about a corresponding P . For this reason, we can assume in what follows that S contains only relation symbols, or only relation symbols and individual constants (which are really O -ary operation symbols), and nevertheless, when we want to, apply our results at once to structures of arbitrary similarity type.

We are now ready to prove some theorems of model theory. However, many of these theorems and their proofs require an (elementary) use of ordinals and cardinals, so we close §1 with a brief discussion of set-theoretical notions.

An ordinal α is the order-type of a well-ordered set. It is convenient to interpret ordinals so that $\alpha = \{\beta: \beta < \alpha\}$. ' ξ ', ' η ', ' α ', ' β ', ' γ ' always denote ordinals. ω is the first infinite ordinal and the set of all natural numbers. The cardinal number \bar{X} of a set X is the smallest ordinal α such that α and X can be put in one-to-one correspondence. (' κ ', ' λ ', ' μ ' always denote cardinals.) The infinite cardinals are written in order as $\omega = \aleph_0, \aleph_1, \aleph_2, \dots$. κ^+ is the first cardinal greater than κ (so $(\aleph_\alpha)^+ = \aleph_{\alpha+1}$). κ is **regular** if it is not the sum of a smaller number of smaller cardinals. Any cardinal of the form $\aleph_{\alpha+1}$ is regular; \aleph_ω is the first cardinal which is not regular.

2. The compactness theorem. We shall now prove the basic theorem of model theory:

2.1 COMPACTNESS THEOREM. *If every finite subset of Σ (a set of S -sentences) has a model, then so has Σ .*

2.1 was proved by Gödel in 1930 for countable S , and extended later by Malcev, Henkin, and A. Robinson to arbitrary S . The languages $L(S)$ with S uncountable might be considered 'imaginary,' but there is no difficulty in considering them mathematically and it turns out that they are very useful as a tool in obtaining results about 'real' languages (with finite or denumerable S).

To save time we agree to say that Σ is **finitely satisfiable** if every finite subset of Σ has a model.

In ordinary mathematics, speaking roughly, when we arrive at the knowledge that $\exists x\phi(x)$, then we usually find it helpful to introduce a name for such an x by saying: 'let a be such an x ', 'choose such an x ', etc. Later on we may do the same thing for another formula $\exists y\psi(y)$, and so on. In the first step toward proving 2.1, we now do something formal analogous to this—in a systematic way covering *all* ϕ . However, in order to avoid having to 'know that $\exists x\phi(x)$ ', we pass to the formula $\exists y(\exists x\phi(x) \rightarrow \phi(y))$, which clearly always holds.

Now speaking precisely, let $\exists u\phi$ be an S sentence, c an individual constant not in S , and \mathfrak{A} an S -structure. Then clearly there is an $a \in A$ such that $(\mathfrak{A}, a) \vdash \exists u\phi \rightarrow \phi(c)$. Thus

(1) Any \mathfrak{A} can be expanded to a model of $\exists u\phi \rightarrow \phi(c)$.

We now want to form many new sentences like $\exists u\phi \rightarrow \phi(c)$ (each with a new c), one corresponding to each ϕ , even including ϕ 's involving new c 's already introduced earlier. One way of doing this precisely is as follows: Let $\kappa = \max(\aleph_0, \bar{S})$. Let c_ξ ($\xi < \kappa$) be distinct constants not in S , and let $S' = S \cup \{c_\xi: \xi < \kappa\}$. By simple cardinal arithmetic there are exactly κ S' -sentences of the form $\exists w\phi$. Let $\exists u_0\phi_0, \dots, \exists u_\xi\phi_\xi, \dots$ ($\xi < \kappa$) be a list of all of them. By recursion, for $\xi < \kappa$, let d_ξ be the first (in the list $c_0, c_1, \dots, c_\eta, \dots$) such that d_ξ differs from d_ζ for all $\zeta < \xi$ and d_ξ does not occur in ϕ_ζ for $\zeta \leq \xi$. (Thus d_ξ is 'new' when introduced.) Now:

(2) Let Δ_S be the set of all the sentences $\exists u_\xi\phi_\xi \rightarrow \phi_\xi(d_\xi^*)$ for $\xi < \kappa$.

By applying (1) repeatedly (with various S 's) we obtain

(3) *Any S -structure can be expanded to a model of any finite subset of Δ_S (and indeed of the whole Δ_S , though we never need this).*

Hence clearly:

(4) *If Σ is a finitely satisfiable set of S -sentences, then $\Sigma \cup \Delta_S$ is also finitely satisfiable.*

We are now ready for the second and last step in proving 2.1. In fact, by (4), 2.1 will be established if we can show,

LEMMA 2.2. *If $\Sigma \cup \Delta_S$ is finitely satisfiable, then it has a model \mathfrak{U} .*

Proof of 2.2. Throughout we are considering $L(S')$. By Zorn's lemma, we can extend $\Sigma \cup \Delta_S$ to a maximal finitely satisfiable set Σ' of sentences. Note that (a) for any σ , either $\sigma \in \Sigma'$ or $\sim\sigma \in \Sigma'$. Indeed, if not, there are finite $\Sigma_1, \Sigma_2 \subseteq \Sigma'$ such that $\Sigma_1 \cup \{\sigma\}$ has no model and $\Sigma_2 \cup \{\sim\sigma\}$ has no model; hence $\Sigma_1 \cup \Sigma_2$ has no model, a contradiction. Secondly, note that (b) if $\sigma_0, \dots, \sigma_{n-1} \in \Sigma'$ and $\{\sigma_0, \dots, \sigma_{n-1}\} \vdash \sigma$ then $\sigma \in \Sigma'$. For if not, then (by (a)) $\sim\sigma \in \Sigma'$ and $\{\sigma_0, \dots, \sigma_{n-1}, \sim\sigma\}$ is a finite subset of Σ' having no model.

Put $C = \{c_\xi: \xi < \kappa\}$. If $c, d \in C$, let cEd if and only if the sentence $c \approx d$ belongs to Σ' . Then E is an equivalence relation on C . Indeed, suppose, for example, that cEd and $c'Ed$. Then $c \approx d$ and $c' \approx d$ belong to Σ' so, by (b), $c \approx c' \in \Sigma'$ and so cEc' . Our model \mathfrak{U} for Σ' will have as its universe A , the set of all equivalence classes $\bar{c} = c/E$ for $c \in C$. The denotation $c^{\mathfrak{U}}$ of each c will be \bar{c} . If $P \in S$, P n -ary, then we put: $P^{\mathfrak{U}}d_1 \dots d_n$ if and only if $Pd_1 \dots d_n \in \Sigma'$; that $P^{\mathfrak{U}}$ is well-defined follows easily again from (b). As remarked already, we can assume that S contains only relation symbols, so \mathfrak{U} is now completely defined.

Finally, we show that (c) for any σ , $\sigma \in \Sigma'$ if and only if $\mathfrak{U} \vdash \sigma$. (Of course, it follows that \mathfrak{U} is a model of $\Sigma \cup \Delta_S$ as desired.) (c) is proved by induction (on the number of occurrences of \exists , \wedge , \sim in σ). The case when σ is atomic has really been dealt with already. The cases when σ is a negation or conjunction, and one direction when σ is $\exists u\phi$ are argued again using (a) and (b) (and inductive hypotheses as below). We illustrate by doing the remaining case. Suppose σ is $\exists u\phi$ and $\sigma \in \Sigma'$. Then we use the fact that, by its construction, Δ_S (and hence Σ') contains some sentence

$\exists u\phi \rightarrow \phi(\frac{u}{c})$. Since also $\exists u\phi \in \Sigma'$, we see from (b) that $\phi(\frac{u}{c}) \in \Sigma'$. So, by the inductive hypothesis, $\mathfrak{U} \vdash \phi(\frac{u}{c})$, whence $\mathfrak{U} \vdash \exists u\phi$ as desired.

For later use, we remark that the proof of 2.2 actually showed a little more, namely:

2.3 In 2.2, \mathfrak{U} can be found such that $A = \{c_\xi^{\mathfrak{U}} : \xi < \kappa\}$.

In the argument above we have not really used the full (b), but only a few much simpler principles, such as: if both σ_1 and $\sigma_1 \rightarrow \sigma_2$ belong to Σ' , then $\sigma_2 \in \Sigma'$; and: every sentence $\phi(\frac{u}{c}) \rightarrow \exists u\phi \in \Sigma'$. If these are systematically formulated and the whole proof of 2.1-recast somewhat, a proof is obtained for Gödel's Completeness Theorem (in fact the proof by Henkin; see e.g., [3]). This famous theorem says, roughly, that all consequences of a set of sentences can be generated by certain natural 'axioms' and 'rules of inferences'—which are thus 'complete' for this purpose. In a considerable part of model theory (and in almost all of this article) compactness is central but completeness does not enter. Therefore, we have given a direct proof of compactness, avoiding the refinements needed to obtain also completeness. A second proof of compactness (also direct) will be given in §5.

It is worth remarking explicitly that an equivalent form of the compactness Theorem 2.1 is:

(5) If $\Sigma \vdash \sigma$ then for some finite $\Sigma' \subseteq \Sigma$, $\Sigma' \vdash \sigma$.

In later sections we shall use the Compactness Theorem several times to obtain various general theorems of model theory. But we could give right now a large number of direct special applications. A typical example is a proof of the existence of a non-Archimedean ordered field $\mathfrak{U} = (A, +, \cdot, <, 1)$. \mathfrak{U} is Archimedean if for any $a \in A$ then exists a positive integer n such $a < n \circ 1$. Let σ be the conjunction of the usual (clearly elementary) axioms for ordered fields. Let c be a new individual constant and let Σ consist of σ together with the (infinitely many) axioms

$$1 < c, 1 + 1 < c, \dots, \overbrace{1 + \dots + 1}^{n \text{ times}} < c, \dots$$

It is clear that for any finite subset Σ' of Σ we can find a real number r such that $(\text{Reals}, +, \cdot, <, 1, r)$ is a model of Σ' . Hence by 2.1 Σ has a model $(A, +, \cdot, <, 1, a)$, and $(A, +, \cdot, <, 1)$ is clearly a non-Archimedean ordered field. By replacing σ in this argument by the (infinite) set of *all* sentences true in $\mathfrak{U}_0 = (\text{Reals}, +, \cdot, <, 1)$ we even get a non-Archimedean ordered field having exactly the same true elementary sentences as the reals \mathfrak{U}_0 .

Non-Archimedean ordered fields can be constructed in other ways. However, these require at least a little special knowledge or ingenuity, while the above argument is extremely simple and general. The reader will have no trouble after seeing it in finding many interesting similar arguments. Incidentally, in these arguments, some-

times no new individual constant is needed, sometimes infinitely many (for example, to create a descending sequence).

Show, for example, that (the property of being a) 'torsion-free group' is not \mathcal{EC} . Or that 'torsion group' is not \mathcal{EC}_Δ and not even \mathcal{PC}_Δ (very easy). Show that 'well-ordered' is not \mathcal{PC}_Δ (very easy). Finally, show that 'not well-ordered' is not \mathcal{EC}_Δ . (Hint: show that the set of all sentences, true in $(\omega, <)$, has a non-well-ordered model.) For harder problems, consider the notions 'free group' and 'simple group'.

3. Löwenheim-Skolem Theorems. Let \mathfrak{A} and \mathfrak{B} be S -structures. \mathfrak{A} and \mathfrak{B} are called **elementarily equivalent**, written $\mathfrak{A} \equiv \mathfrak{B}$, if they have the same true elementary sentences. For an easy but instructive example, consider the three non-isomorphic groups (Integers, +), (Rationals, +), and (Non-zero reals, \cdot), and verify that no two of these are elementarily equivalent.

We have just seen that there is an ordered field which is elementarily equivalent to the reals \mathfrak{U}_0 but, being non-Archimedean, is certainly not isomorphic to \mathfrak{U}_0 . Thus, in general, \equiv is definitely weaker than \cong . However, if \mathfrak{A} (i.e., A) is finite and also S is finite, it is trivial to write down a single sentence σ true in \mathfrak{A} such that any model of σ is isomorphic to \mathfrak{A} . Even if S is infinite, but A is still finite, every \mathfrak{B} elementarily equivalent to \mathfrak{A} is isomorphic to \mathfrak{A} . (Inferring this from the (preceding) result for finite S is amusing, though not difficult.) On the other hand, we shall see in 3.3 below, that for *every* infinite \mathfrak{A} , this fails—indeed, for each infinite $\kappa (\geq \bar{S})$, there exists a structure \mathfrak{B} of power κ which is elementarily equivalent to \mathfrak{A} .

There is another very useful notion related to but stronger than elementary equivalence. As always, \mathfrak{A} is called a **substructure** of \mathfrak{B} if $A \subseteq B$, A is closed under the (distinguished) operations of \mathfrak{B} , and each relation or operation of \mathfrak{A} is the natural restriction to A of the corresponding relation or operation of \mathfrak{B} . We write $\mathfrak{A} = \mathfrak{B} \upharpoonright A$. \mathfrak{A} is called an **elementary substructure** of \mathfrak{B} (written $\mathfrak{A} < \mathfrak{B}$) if in addition: for any n , any $a_0, \dots, a_{n-1} \in A$ and any n -formula θ , $\mathfrak{A} \models \theta[a_0, \dots, a_{n-1}]$ if and only if $\mathfrak{B} \models \theta[a_0, \dots, a_{n-1}]$. Clearly $\mathfrak{A} < \mathfrak{B}$ means that \mathfrak{A} is a substructure of \mathfrak{B} and $\mathfrak{A} \equiv \mathfrak{B}$, but it means more too. Thus, $(\omega - \{0\}, <)$ is even isomorphic to $(\omega, <)$ but is not an elementary substructure (because 'is the first element' can be expressed by an elementary formula).

LEMMA 3.1. *Let \mathfrak{A} be a substructure of \mathfrak{B} . Suppose (*) for any $a_0, \dots, a_{n-1} \in A$, any $b \in B$, and any $(n+1)$ -formula ϕ , if $\mathfrak{B} \models \phi[a_0, \dots, a_{n-1}, b]$ then for some $a_n \in A$, $\mathfrak{B} \models \phi[a_0, \dots, a_{n-1}, a_n]$. Then $\mathfrak{A} < \mathfrak{B}$.*

The reader will easily establish 3.1 by induction on the formula θ (in the definition of $<$).

The converse of 3.1 is almost immediate. Thus (*) gives a characterization of $\mathfrak{A} < \mathfrak{B}$ which refers only to satisfaction in \mathfrak{B} . Indeed, (*) can be regarded as saying (as in algebra) that A is a closed subset of \mathfrak{B} in a certain sense (with respect to all ϕ).

The next theorem is perhaps the oldest theorem in model theory, and one of the most fundamental. The first version appeared in 1915.

THEOREM 3.2 (Downward Löwenheim-Skolem Theorem). *Suppose \mathfrak{B} is an S -structure, κ is an infinite cardinal, and $\bar{S} \leq \kappa \leq \bar{B}$. Then \mathfrak{B} has an elementary substructure \mathfrak{A} of power κ .*

The proof is little more than the proof, say, that any countable subset of a group generates a countable subgroup. Choose a subset X of B of power κ . Now well-order B . Then each, say, $(n+1)$ -formula ϕ determines an n -ary (partial) operation taking any b_0, \dots, b_{n-1} into the 'first' b such that $\mathfrak{B} \models \phi[b_0, \dots, b_{n-1}, b]$. By 3.1, if we take for A the closure of X under all these operations, then $\mathfrak{A}(=\mathfrak{B} \upharpoonright A) < \mathfrak{B}$. Since there are at most κ of our operations, clearly $\bar{A} = \kappa$.

If, in addition to the hypothesis of 3.2, a subset Y of A of power $\leq \kappa$ is given, then \mathfrak{B} can also be required to include Y . This is clear from the proof, or can be inferred by applying 3.2 to the structure $(\mathfrak{A}, y)_{y \in Y}$.

In the simplest case, 3.2 says that any infinite \mathfrak{B} (whose S is countable) has a countable elementary substructure \mathfrak{A} . If we take for \mathfrak{B} the class of all sets together with the ϵ -relation, then we obtain a denumerable 'world' \mathfrak{A} of sets in which all the axioms of transfinite set theory hold — a state of affairs sometimes called Skolem's 'paradox'. Though there is apparently no real paradox, the existence of such an \mathfrak{A} is of basic importance in the remarkable work of Gödel and Cohen on the consistency and independence of the continuum hypothesis and other basic propositions in set theory.

It is very significant that in 3.2 the smaller structure \mathfrak{A} which is obtained is not only elementarily equivalent to \mathfrak{B} but is also a **substructure** of \mathfrak{B} (and even an elementary substructure). Suppose, for example, that $\mathfrak{B} = (B, <)$, where $<$ well-orders B . As we have seen, a structure can be elementarily equivalent to \mathfrak{B} but not well-ordered. But obviously any substructure of a well-ordering is a well-ordering, so 3.2 gives a countable well-ordering \mathfrak{A} elementarily equivalent to \mathfrak{B} . (And in the set theory example, the 'ordinals' of the countable 'world' are really well-ordered.)

Another important fact is that one can always add relations to \mathfrak{B} before applying 3.2. A good example of this will be seen later on in the beginning of the proof of 7.1 (where $<$ is adjoined to \mathfrak{A}). We shall next prove a strengthening 3.2' of 3.2 in which the elementary language L is replaced by a richer language. It will be quicker (and also instructive) to prove 3.2' anew by imitating the proof of 3.2. However, we could instead infer 3.2' directly from 3.2, by systematically employing the device of adding new relations to \mathfrak{B} before applying 3.2. (Lemma 7.5 in §7 tells just what relations to add.)

We now consider for the first time a language richer than the elementary language. The language $L_{\omega_1, \omega}(S)$ has all the symbols of $L(S)$ and in addition the symbols \vee and \wedge , which are used in forming countably *infinite* disjunctions and conjunctions.

(Thus this language is certainly ‘imaginary’, as a single formula can be infinitely long.) The definition of **formula** (of $L_{\omega_1\omega}(S)$) is obtained from the old definition by adding the clause: if, for some k , $\langle \phi_0, \dots, \phi_n, \dots \rangle$ is an arbitrary sequence of k -formulas, then $\bigvee (\phi_0, \dots, \phi_n, \dots)$ and $\bigwedge (\phi_0, \dots, \phi_n, \dots)$ are formulas. (Strictly speaking, this means that the set of formulas is defined as the smallest set K meeting the old conditions and closed under the formation of disjunctions and conjunctions as above. For convenience we often write, say, the disjunction $\bigvee (\phi_0, \dots, \phi_n, \dots)$ as $\bigvee_n \phi_n$ or as $\phi_0 \vee \phi_1 \vee \phi_2 \vee \dots$.)

As an example, we can say in a single sentence of $L_{\omega_1\omega}$ that a group $\mathcal{G} = (G, \circ, e)$ is a torsion group. The sentence is:

$$\forall u(u \approx e \vee u \circ u \approx e \vee u \circ u \circ u \approx e \vee \dots).$$

The clearly intended meaning for satisfaction and truth is made precise just as before and the notation \vdash is retained. Notice that a restriction was made in defining ‘formula’ so that every formula (of $L_{\omega_1\omega}$) has only finitely many free variables. (Those left out could never have been parts of sentences anyway.)

Even if S is finite, $L_{\omega_1\omega}(S)$ has uncountably many formulas, so we usually must content ourselves with looking at ‘fragments’ (i.e., subsets) of the set of all formulas (which can already be extremely rich). In particular, a formula ϕ determines a very important fragment, the set of all **subformulas** of ϕ . This notion has an obvious meaning, but formally, the definition is inductive: The subformulas of $\sim \phi$ are $\sim \phi$ and those of ϕ ; the subformulas of $\bigvee_n \phi_n$ are $\bigvee_n \phi_n$ and all subformulas of any ϕ_n ; etc. By induction (based on the inductive definition of ‘formula’), the set of subformulas of any formula ϕ is immediately seen to be countable. Now let Φ be any set of formulas of $L_{\omega_1\omega}(S)$. We write $\mathfrak{A} \prec (\Phi) \mathfrak{B}$ to mean that for any $\phi \in \Phi$ and any $a_0, \dots, a_{n-1} \in A$, $\mathfrak{A} \vdash \phi[a_0, \dots, a_{n-1}]$ if and only if $\mathfrak{B} \vdash \phi[a_0, \dots, a_{n-1}]$. In place of 3.1 we have:

LEMMA 3.1'. $\mathfrak{A} \prec (\theta) \mathfrak{B}$ if (*) of 3.1 holds with ‘any ϕ ’ replaced by ‘any subformula ϕ of θ ’.

3.1' is proved just as was 3.1; one shows that $\mathfrak{A} \prec (\psi) \mathfrak{B}$ for all subformulas ψ of θ , by induction on ψ . (Clearly 3.1' slightly improves 3.1 even for the elementary language.)

Arguing exactly as in the proof of 3.2 (but using 3.1') one obtains the downward Löwenheim-Skolem theorem for $L_{\omega_1\omega}$:

THEOREM 3.2'. Suppose \mathfrak{B} is an S -structure, κ is an infinite cardinal, $\kappa \leq \bar{B}$, and Φ is a set of $L_{\omega_1\omega}(S)$ -formulas of power at most κ . Then there exists a structure \mathfrak{A} of power κ such that $\mathfrak{A} \prec (\Phi) \mathfrak{B}$.

In particular, for a single formula θ of $L_{\omega_1\omega}$, there exists a countable $\mathfrak{A} \prec (\theta) \mathfrak{B}$.

Examples of the use of 3.2 and 3.2' will be given a little later. We first return to the elementary language and prove the 'upward' Löwenheim-Skolem theorem (which is actually due to Tarski). This result, unlike 3.2, will be obtained by means of the compactness theorem.

THEOREM 3.3 (Upward Löwenheim-Skolem Theorem). *Suppose the set Σ of S -sentences has an infinite model \mathfrak{B} . Then Σ has a model in each infinite power $\kappa \geq \bar{S}$.*

Of course an equivalent wording (parallel to 3.2), obtained by taking Σ to be the set of all sentences true in \mathfrak{B} , says that in each such power there exists $\mathfrak{U} \equiv \mathfrak{B}$.

Proof. Let d_0, \dots, d_ξ, \dots ($\xi < \kappa$) be distinct new constants and let Σ' be the set of all sentences $d_\xi \neq d_\eta$ ($\xi < \eta < \kappa$). Clearly every finite subset of $\Sigma \cup \Sigma'$ has a model, obtained by adjoining individuals to the given infinite model. (In fact it would clearly be enough to know that Σ has arbitrarily large finite models.) Hence by the compactness theorem, $\Sigma \cup \Sigma'$ has a model \mathfrak{U} . Clearly $\bar{A} \geq \kappa$. Since $\bar{S} \leq \kappa$, we could, by 2.3, have taken \mathfrak{U} so that also $\bar{A} \leq \kappa$. (Or, we could take any \mathfrak{U} and apply the Downward Löwenheim-Skolem theorem to obtain another of power κ .)

Observe that 3.3 is important when S is finite, and even in this case the proof makes a detour to a language which has κ symbols.

3.2 and 3.3 overlap. For example, they both show that (a) if S is countable and \mathfrak{B} infinite then there is a countable $\mathfrak{U} \equiv \mathfrak{B}$. 3.2, however, gives \mathfrak{U} as a substructure of \mathfrak{B} , which as we have seen is much stronger. (On the other hand, using the ideas of the proofs of 3.3 and 2.1, one can establish (a) without using the axiom of choice, which is involved essentially in the proof of 3.2.)

The Löwenheim-Skolem theorems have many applications. In order to discuss some of them, we first introduce some simple and natural terminology. A set T of S -sentences is called a **theory** if for any S -sentence σ , if $T \vdash \sigma$ then $\sigma \in T$. Clearly a theory T (unlike an arbitrary Σ) specifies its S , called $S(T)$. Any class K of S -structures determines a theory called $Th(K)$, namely, the set of all S -sentences true in all members of K . For example, we can speak of the theory of groups (i.e., $Th(K)$, where K is the class of all groups). For single structures, clearly $\mathfrak{U} \equiv \mathfrak{B}$ if and only if $Th(\mathfrak{U}) = Th(\mathfrak{B})$. A set Δ of S -sentences is called a set of **axioms** for the theory T if T is just the set of all consequences of Δ . We denote by $T + \Sigma$ the theory T' axiomatized by $T \cup \Sigma$ and such that $S(T')$ is $S(T)$ plus the non-logical symbols occurring in Σ . A theory T is called **complete** (a usage quite different from that in 'completeness theorem') if T has a model and for any $S(T)$ -sentence σ , either $\sigma \in T$ or $\sim \sigma \in T$; or equivalently, if for some \mathfrak{U} , $T = Th(\mathfrak{U})$.

A beautiful example of the application of 3.2' and 3.3 is given by the theory T_0 of algebraically closed fields of characteristic zero. As is familiar (see, e.g., [16]), these are fields $\mathfrak{U} = (A, +, \cdot)$ in which (1) $n \circ 1 \neq 0$ for any positive integer n , and (2) every polynomial (of degree ≥ 1) has a root. The field of complex numbers

is such a field and so is the field of algebraic (complex) numbers. Though (1) and (2) sound fairly complicated, it is trivial to see that the class K_0 of all algebraically closed fields of characteristic zero is elementary in the wider sense (EC_A). Indeed, a few well-known elementary sentences say that \mathfrak{U} is a field; (1) can be replaced by the infinite set of sentences:

$$(1') \quad 1 + 1 \neq 0, 1 + 1 + 1 \neq 0, \dots;$$

and (2) can be replaced by all the sentences

$$(2') \quad (\forall u_0, u_1, \dots, u_n) [u_0 \neq 0 \rightarrow \exists x (u_0 x^n + u_1 x^{n-1} + \dots + u_n = 0)]$$

$$(n = 1, 2, 3, \dots).$$

(Here 0 and 1 are elementarily 'definable' from $+$ and \cdot and could be eliminated; while entries like x^k , say for $k = 6$, should be replaced by $xxxxxx$.)

Recall that in a field \mathfrak{U} an element x is said to be (algebraically) dependent on the elements y_0, \dots, y_{m-1} if some equation $u_0 x^m + u_1 x^{m-1} + \dots + u_n = 0$ holds in which the u_i , not all 0, are obtainable as polynomials in y_0, \dots, y_{m-1} with integral coefficients. A subset X of A is called independent if no element of X is dependent on finitely many other elements of X . A maximal independent subset of A is called a basis for \mathfrak{U} . The following facts (see e.g., [16]), go back to Steinitz: (a) Any two bases for \mathfrak{U} have the same cardinal number, called the degree of transcendence of \mathfrak{U} . (b) For each cardinal κ there is (up to isomorphism) exactly one algebraically closed field of characteristic zero and transcendence degree κ .

If the transcendence degree of \mathfrak{U} is $0, 1, \dots, n, \dots$ or \aleph_0 then clearly \mathfrak{U} is denumerable (and all these \mathfrak{U} 's are nonisomorphic). On the other hand if the transcendence degree of \mathfrak{U} is an uncountable cardinal κ , then already it coincides with \bar{A} . Hence, for each uncountable cardinal κ , there is exactly one model of T_0 of power κ . In general, this situation is described by saying that T_0 is **categorical in the power κ** . Theorem 3.3 has an immediate but remarkable consequence regarding such theories.

COROLLARY 3.4. *If a countable theory T is categorical in some infinite power κ and has no finite models, then T is complete.*

Indeed, any model \mathfrak{U} of T is elementarily equivalent to a model \mathfrak{B} of power κ , and \mathfrak{B} is unique (up to isomorphism).

Apply 3.4 to T_0 , we obtain:

THEOREM 3.5. *Any two algebraically closed fields of characteristic zero are elementarily equivalent.*

Thus, if one establishes an **elementary** statement about the complex number field by a detour through complex analysis, he can nevertheless be sure that the statement holds also in the field of algebraic numbers, and indeed, in any algebraically closed field of characteristic zero!

3.5 was first established by another method by Tarski [15]. The above method (or others) can be pushed further to yield information about what sets and relations are elementarily definable in a model of T_0 (cf. [1], [15]).

Now we consider what can be said about fields $\mathfrak{A} \in K_0$ in the language $L_{\omega_1\omega}$. It is easy to find $L_{\omega_1\omega}$ -sentences $\sigma_0, \dots, \sigma_n, \dots, \sigma_\omega$ such that σ_n (or σ_ω) says that the transcendence degree is n (or, is infinite). (By 3.5, no such *elementary* sentences can be found.) Verifying this (after rereading the example above of a sentence of $L_{\omega_1\omega}$ saying 'is a torsion group') will give the reader some good experience with $L_{\omega_1\omega}$.

On the other hand, by 3.2', if $\mathfrak{B} \in K_0$, \mathfrak{B} is uncountable, and ϕ is any single sentence of $L_{\omega_1\omega}$ holding in \mathfrak{B} , then there is a countable $\mathfrak{A} \in K_0$ of transcendence degree \aleph_0 such that \mathfrak{A} is a model of ϕ . (For Φ in 3.2' take $\{\phi, \sigma_\omega, \sigma\}$, where σ is the conjunction of all the elementary axioms characterizing K_0). But there is up to isomorphism only one such \mathfrak{A} , so we conclude:

THEOREM 3.6. *Any two algebraically closed fields of characteristic zero having infinite transcendence degrees are $L_{\omega_1\omega}$ -equivalent (i.e., have the same true $L_{\omega_1\omega}$ -sentences).*

The fact that nearly anything you could prove about the complex number field apparently could always later be proved for all algebraically closed fields of characteristic zero was observed by algebraists and algebraic geometers (and sometimes called "Lefschetz' Principle"). It is interesting that when made precise and proved, this conjecture splits into the overlapping statements 3.5 and 3.6.

We have now completed the mathematical development of this section and are ready for the next. But it is impossible not to make several informal remarks about some very important topics closely related to the discussion above.

Firstly: when (as in 3.5) the completeness of a theory $T(=T_0)$ is obtained, usually the so-called **decidability** of T is also obtained. Roughly speaking, T (or even an arbitrary set Σ of S -sentences), is called **decidable** or **recursive** if a machine can be built which can decide, given any S -sentence, whether or not $T \vdash \sigma$ (or $\sigma \in \Sigma$). (A correct discussion requires the theory of recursive sets and functions.) In general, if T has a recursive set of axioms, as is usual in practice, then (the hypothesis of) 3.4 does yield the decidability of T , but only in a useless, highly theoretical sense. The proof of 3.5 by Tarski, however, gives a stronger decidability for T_0 , implying that a 'reasonable' machine might be built. As regards decidability, the best that can be said for results like 3.4, is that they may, in some cases, alert us to try to find a reasonable decision procedure.

Secondly: Although 3.3 is only concerned with the existence of structures elementarily equivalent to a given \mathfrak{A} , we used it to show the elementary equivalence of given structures \mathfrak{A} and \mathfrak{B} , for example, the field of complex numbers and the field of all algebraic complex numbers. But theories categorical in power are rare,

and in general, other methods must be used to prove elementary equivalence (or $L_{\omega_1\omega}$ -equivalence) of given structures. Examples of such methods are the method of eliminating quantifiers (cf. [15]), the method of model-completeness (cf. [12]), and the Fraïssé-Ehrenfeucht method (cf. [3]). Among the four or five most famous results (on completeness and decidability) are Tarski's [15] concerning the field of real numbers and that of Ax and Kochen [1] concerning the p -adic number fields. There is a great deal still to be done. For example, it is an open problem (raised long ago by Tarski) whether the free group with two generators and the free group with three generators are elementarily equivalent!

Finally, let us look more closely at theories T categorical in some infinite power. The theory T_0 of algebraically closed fields of characteristic zero is categorical in each power $\kappa > \aleph_0$ but not in the power \aleph_0 . The theory of Abelian groups in which every element is of order p (p a fixed prime) is easily seen to be categorical in every infinite power (again by a 'basis' argument, now a vector space basis with respect to the field $\{0, \dots, p-1\}$). The theory of densely ordered sets $(A, <)$ with no extreme elements is, as is well known (cf. §4), categorical in the power \aleph_0 but no other. It was conjectured by Łoś that these three behaviours are the only possibilities, and Morley gave several years later a beautiful proof of this conjecture:

THEOREM 3.7 [9]. *If a countable theory T is categorical in one uncountable power, then it is categorical in every uncountable power.*

We do not prove 3.7, and indeed present it partly just as an example of a simple, intuitively motivated statement of model theory which requires a deep proof. The reader may find the original proof in [9] and an interesting recent proof in a paper by Baldwin and Lachlan [18]. (To read these, nothing beyond what is in this article is required.)

There are a number of further conjectures regarding theories categorical in uncountable powers which the evidence suggests (in particular, think of T_0). With a good deal of further effort and ingenuity, several of them have been established. Thus, Morley [10] showed that for such a T the number of nonisomorphic denumerable models must be countable. A general result (see 6.7) tells us that (for any complete theory) the number cannot be 2, but Morley's work left open whether it might be 3, 4, 5, ... (as well as 1 and \aleph_0 as in examples above). Then Baldwin and Lachlan (in the aforementioned article) showed that it must be 1 or \aleph_0 . Another problem on which several authors made contributions was that of extending Morley's result 3.7 to arbitrary T ("uncountable power" being replaced by "power $> \bar{T}$ "). This has only very recently been established in full generality, by S. Shelah (in a paper to appear in the Proceedings of the Tarski Symposium held in Berkeley, 1971).

Perhaps the most interesting questions still unresolved are those concerning the finite axiomatizability of T : (a) if T is categorical in \aleph_1 and complete (i.e., by 3.4, has no finite models), can T be finitely axiomatizable? (b) If T (possibly incom-

plete) is categorical in \aleph_1 but not in \aleph_0 , can T be finitely axiomatizable? (These two problems were formulated by J. Ax as a refinement of one problem in [9], p. 537.)

4. Saturated models. Before beginning on the main topic of this section, we must call attention to a second basic fact about the notion $<$, namely 4.1 below, which is as pervasive in model theory as the downward Löwenheim-Skolem theorem (3.2). Let S be a set of relation symbols. Suppose \mathcal{C} is a set of S -structures directed by $<$, i.e., such that if $\mathfrak{A}, \mathfrak{B} \in \mathcal{C}$, then for some $\mathfrak{C} \in \mathcal{C}$, $\mathfrak{A} < \mathfrak{C}$ and $\mathfrak{B} < \mathfrak{C}$. Then (as in algebra) we define $\bigcup(\mathfrak{A}: \mathfrak{A} \in \mathcal{C})$ to be the S -structure \mathfrak{B} such that $|\mathfrak{B}| = \bigcup(|\mathfrak{A}|: \mathfrak{A} \in \mathcal{C})$ and for each $P \in S$, $P^{\mathfrak{B}} = \bigcup(P^{\mathfrak{A}}: \mathfrak{A} \in \mathcal{C})$.

THEOREM 4.1. (Tarski) *If \mathfrak{B} is the union of a family \mathcal{C} of S -structures directed by $<$, then for each $\mathfrak{A} \in \mathcal{C}$, $\mathfrak{A} < \mathfrak{B}$.*

4.1 is easily proved by induction on the length of the formula θ in the definition of ' $\mathfrak{A} < \mathfrak{B}$ '. (Like 3.2, it can be extended in suitable forms to $L_{\omega_1, \omega}$ —but we shall not need this.)

Cantor established a famous theorem about the ordering $(Q, <)$ of the rational numbers:

- (1) Any two denumerable dense orderings $(A, <)$ and $(A', <')$ without end points are isomorphic;
- (2) Any countable ordering is isomorphic to a substructure of $(Q, <)$;
- (3) If $q_0 < q_1 < \dots < q_n$ and $q'_0 < q'_1 < \dots < q'_n$, then there is an automorphism of $(Q, <)$ taking q_i into q'_i ($i \leq n$).

Since this whole section will be a generalization of Cantor's theorem, we briefly recall the proof.

For (1), we first write $A = \{a_0, \dots, a_n, \dots\}$ and $A' = \{a'_0, \dots, a'_n, \dots\}$. The desired isomorphism f will be $\{(x_n, x'_n): n \in \omega\}$, where the (x_n, x'_n) are defined by recursion on n in such a way that $(*)x_j < x_k$ if and only if $x'_j < x'_k$. Suppose we already have (x_i, x'_i) , for $i < n$, such that $(*)$ holds (for $j, k < n$). If n is even, let x_n be the first unused a_i (in the list a_0, a_1, \dots). The hypotheses insure that there is an a'_j having exactly the same order-relations with x'_0, \dots, x'_{n-1} that x_n has with x_0, \dots, x_{n-1} ; we take such an a'_j for x'_n . If n is odd, proceed symmetrically; i.e., let x'_n be the first unused a'_i and choose x_n to match its order relations. The f so obtained is clearly an isomorphism of \mathfrak{A} onto \mathfrak{A}' . The proof of (3) is similar, and that of (2) simpler, as the 'back and forth' procedure (distinguishing even and odd n) is omitted.

After seeing the method of Cantor's proof, anyone is likely to speculate that the method can be used in other places. For instance, one could ask for a dense ordering in other powers $> \aleph_0$ with properties similar to (1), (2), and (3) (cf. the η_α -orders of Hausdorff). Generalizing more broadly, we could raise similar questions with any other class K of similar structures such as the class of all groups in place of the class of orderings. The strongest results of this type were obtained by Jonsson [4]. Of course, they require special hypotheses about the class K . However, it turns

out that if algebraic notions like substructure (in (2) and in Jonsson's work) are replaced by model-theoretic notions like elementary substructure, then some useful and quite uniform results of general model theory are obtained. We now establish these results (due to Morley and the author [11]). (We give a direct argument, but the results can also be inferred from Jonsson's broader, algebraic results by a suitable interpretation.)

Roughly speaking, the role of the class of orderings is assumed in our work by the class of all models of an arbitrary complete theory T having infinite models. However, since such a T is determined by any model \mathfrak{A} of T , there is sometimes no explicit mention of a theory T .

If $Y \subseteq B$, $f: Y \rightarrow A$, and $(\mathfrak{B}, y)_{y \in Y} \equiv (\mathfrak{A}, f(y))_{y \in Y}$, then we say that f **embeds** Y (over \mathfrak{B}) **elementarily in** \mathfrak{A} . (In context, 'over \mathfrak{B} ' is usually omitted.)

The following usage is nearly self-explanatory: A set Ψ of 1-formulas (of $L(S(\mathbb{C}))$) is **satisfiable in** \mathbb{C} if for some $x \in C$, $\mathbb{C} \vdash \psi[x]$ for all $\psi \in \Psi$. Ψ is **finitely satisfiable in** \mathbb{C} if each finite subset of Ψ is satisfiable in \mathbb{C} .

DEFINITION 4.2. Let \mathfrak{A} be a structure of infinite power κ . We say that:

- (a) \mathfrak{A} is *homogeneous* if any elementary embedding of a subset of A of power $< \kappa$ into \mathfrak{A} can be extended to an automorphism of \mathfrak{A} .
- (b) \mathfrak{A} is *universal* (or μ -*universal*) if, for any $\mathfrak{B} \equiv \mathfrak{A}$ and any subset Y of \mathfrak{B} of power $\leq \kappa$ (or, of power $< \mu$), Y can be elementarily embedded in \mathfrak{A} .
- (c) \mathfrak{A} is *saturated* if whenever X is a subset of A of power $< \kappa$, and Φ is any set of 1-formulas which is finitely satisfiable in $(\mathfrak{A}, x)_{x \in X}$, then Φ is satisfiable in $(\mathfrak{A}, x)_{x \in X}$.

(c) suggests that saturated models might well be called 'compact' models and, in fact, this name has been used for this notion, but also for some related but not equivalent notions.

(c) can be stated in very slightly different, equivalent form involving the useful notion of a type of element. Given a complete S -theory T , choose a new constant c (not in $S(T)$). If \mathfrak{A} is a model of T and $a \in A$, then $Th((\mathfrak{A}, a))$ is a complete $S \cup \{c\}$ -theory, called the **type of a (in \mathfrak{A})**. The set of all such $Th((\mathfrak{A}, a))$, over **all models** \mathfrak{A} of T and elements a of A , is called the set of all **types of elements over T** . In other words, these types are just the complete $S \cup \{c\}$ -theories extending T . If \mathfrak{A} is a model of T and T' a type of element over T , then in general there may or may not be an element a in \mathfrak{A} of type T' ; if there is, we say that T' is realized in \mathfrak{A} . For example let \mathfrak{B} be an ordered field elementarily equivalent to the reals \mathfrak{A}_0 but possessing an element b , such that for each n , $n \circ 1 < b$. Then $Th((\mathfrak{B}, b))$ is a type over \mathfrak{A}_0 (i.e., over $Th(\mathfrak{A}_0)$), but is not realized in \mathfrak{A}_0 . However, a saturated model realizes all possible types (hence the name) and this is so even after not too many 'parameters' are added. Speaking precisely:

4.2(c') \mathfrak{A} is saturated if and only if, for any subset X of A of power $< \bar{A}$, $(\mathfrak{A}, x)_{x \in X}$ realizes all possible types of elements (over itself).

4.2(c') is trivial to verify, using the Compactness Theorem.

Another simple but useful fact about saturated models, immediate using (c), is:

(4) If \mathfrak{M} is saturated and $S' \subseteq S(\mathfrak{M})$ then $\mathfrak{M} \upharpoonright S'$ is saturated.

The next theorem, 4.3, is similar to Cantor's theorems (1), (2), and (3).

THEOREM 4.3. *Let $\bar{\mathfrak{M}} = \kappa \geq \aleph_0$.*

(a) *If \mathfrak{M} and \mathfrak{M}' are both saturated and of power κ and $\mathfrak{M} \equiv \mathfrak{M}'$ then $\mathfrak{M} \cong \mathfrak{M}'$.*

(b) *\mathfrak{M} is homogeneous if (and only if) whenever X is a subset of A of power $< \kappa$, f is an elementary embedding of X into \mathfrak{M} , and $a \in A$, then f can be extended to an elementary embedding of $X \cup \{a\}$ into \mathfrak{M} . (This condition might be called "innerly saturated.")*

(c) *If \mathfrak{M} is saturated then \mathfrak{M} is homogeneous and universal.*

(d) *If \mathfrak{M} is homogeneous and ω -universal then \mathfrak{M} is saturated.*

Proof. (a) We imitate exactly Cantor's proof above of (1). Let $A = \{a_\xi: \xi < \kappa\}$ and $A' = \{a'_\xi: \xi < \kappa\}$. The desired isomorphism f will be $\{(x_\xi, x'_\xi): \xi < \kappa\}$, where the pairs (x_ξ, x'_ξ) are defined by recursion on ξ in such a way that $(*) (\mathfrak{M}, x_\eta)_{\eta \leq \xi} \equiv (\mathfrak{M}', x'_\eta)_{\eta \leq \xi}$. Suppose we are to define (x_ζ, x'_ζ) . Since $(*)$ holds for each $\xi < \zeta$, it is clear that even when ζ is a limit ordinal, $(\mathfrak{M}, x_\eta)_{\eta < \zeta} \equiv (\mathfrak{M}', x'_\eta)_{\eta < \zeta}$. Suppose, for example, ζ is even. Let x_ζ be the first unused a_ξ . 4.2(c') tells us directly that there exists $z \in A'$ such that $(\mathfrak{M}, x_\eta)_{\eta \leq \zeta} \equiv (\mathfrak{M}', x'_\eta, z)_{\eta < \zeta}$; so we can take such a z for x'_ζ .

(b) is proved in a completely similar way. As regards (c), if \mathfrak{M} is saturated, \mathfrak{M} is certainly homogeneous as the right-hand condition in (b) is immediate. To prove \mathfrak{M} is universal (cf. Cantor's (2)), one proceeds just as in the proof of (a), but with no 'back and forth'.

The argument for (d) is amusing. Call \mathfrak{M} λ -saturated if $(\mathfrak{M}, x)_{x \in X}$ realizes all types of elements whenever $|X| < \lambda$. The Cantor-type argument for (c) really shows that (i) if \mathfrak{M} is λ -saturated, then \mathfrak{M} is λ^+ -universal. On the other hand, it is trivial that (ii) if \mathfrak{M} is homogeneous and λ -universal then \mathfrak{M} is λ -saturated, provided λ is infinite and $\lambda \leq \kappa$. (Indeed, suppose $\mu < \lambda$, and T' is a type over $(\mathfrak{M}, x_\xi)_{\xi < \mu}$, say $T' = Th(\mathfrak{B}, y_\xi, b)_{\xi < \mu}$. By λ -universality, we obtain $(\mathfrak{M}, z_\xi, w)_{\xi < \mu} \equiv (\mathfrak{B}, y_\xi, b)_{\xi < \mu}$. By homogeneity, there exists $a \in A$ such that $(\mathfrak{M}, x_\xi, a)_{\xi < \mu} \equiv (\mathfrak{M}, z_\xi, w)_{\xi < \mu}$. Clearly the type of a in $(\mathfrak{M}, x_\xi)_{\xi < \mu}$ is T' .)

By (i) and (ii), if $\aleph_0 \leq \lambda \leq \kappa$ and \mathfrak{M} is homogeneous, then if \mathfrak{M} is λ -universal, \mathfrak{M} is λ^+ -universal. Clearly for a limit cardinal λ , \mathfrak{M} is λ -universal if it is μ -universal for all $\mu < \lambda$. Hence (by induction on the cardinal λ), if \mathfrak{M} is homogeneous and ω -universal, then \mathfrak{M} is κ^+ -universal. It follows trivially (by (ii)) that such an \mathfrak{M} is saturated.

Theorem 4.3 is all very nice but it does not say that there are any saturated structures. The next theorem says that they exist very generally provided we assume the generalized continuum hypothesis (G.C.H.): $2^{\aleph_\alpha} = \aleph_{\alpha+1}$ for all α .

THEOREM 4.4. (G.C.H.) *Given any complete S -theory T , having infinite models,*

and any uncountable, regular power $\kappa > \bar{S}$, there exists a saturated model \mathfrak{U} of T of power κ (and up to isomorphism just one, by 4.3).

Proof. We build up the desired \mathfrak{U} by applying the compactness theorem many times. By 3.3, there is a model \mathfrak{U}_0 of T of power κ . I claim that there exists \mathfrak{U}_1 such that

- (5) $\mathfrak{U}_1 \succ \mathfrak{U}_0$, $\bar{A}_1 = \kappa$, and, for any subset X of A_0 of power $< \kappa$, $(\mathfrak{U}, x)_{x \in X}$ realizes all possible types of elements.

Assume for a minute this claim. Then we can also obtain \mathfrak{U}_2 which is to \mathfrak{U}_1 just as \mathfrak{U}_1 was to \mathfrak{U}_0 . Continuing, we obtain $\mathfrak{U}_0, \mathfrak{U}_1, \dots, \mathfrak{U}_\xi, \dots (\xi < \kappa)$, where, for a limit ordinal ξ , $\mathfrak{U}_\xi = \cup (\mathfrak{U}_\zeta : \zeta < \xi)$. The desired \mathfrak{U} is $\cup (\mathfrak{U}_\xi : \xi < \kappa)$. Indeed, clearly $\bar{A} = \kappa$ and by 4.1, each $\mathfrak{U}_\xi < \mathfrak{U}$. If $X \subseteq A$ and $\bar{X} < \kappa$ then, because κ is assumed regular, X must be a subset of some A_ξ . But any type of element over $Th(\mathfrak{U}, x)_{x \in X} = Th(\mathfrak{U}_\xi, x)_{x \in X}$ is realized in $\mathfrak{U}_{\xi+1}$ by construction (and hence in \mathfrak{U} , since $\mathfrak{U}_{\xi+1} < \mathfrak{U}$). Thus \mathfrak{U} is saturated (by 4.2(c')).

Now we must create a structure \mathfrak{U}_1 for which (5) holds. Consider all pairs (X, T') such that X is a subset of A_0 of power $< \kappa$ and T' is a type of element over $(\mathfrak{U}_0, x)_{x \in X}$. We are assuming κ is regular, $\bar{S} + \aleph_0 < \kappa$, and G.C.H. Using these, it is routine to compute that there are at most κ such X 's and, indeed, at most κ such pairs (X, T') . Suppose we can show that for any one such pair (X, T') there always exists $\mathfrak{U}^{(1)}$ such that

- (6) $\mathfrak{U}_0 < \mathfrak{U}^{(1)}$ and T' is realized in $(\mathfrak{U}^{(1)}, x)_{x \in X}$. Clearly (by 3.2), we can also insist that $\bar{A}^{(1)} \leq \kappa$. Then we can form $\mathfrak{U}_0, \mathfrak{U}^{(1)}, \mathfrak{U}^{(2)}, \dots, \mathfrak{U}^{(\xi)}, \dots (\xi < \kappa)$, one dealing with each pair (X, T') , and take their union for the desired \mathfrak{U}_1 .

Finally, given one pair (X, T') , to obtain (6), put $\bar{\mathfrak{U}} = (\mathfrak{U}_0, a)_{a \in A}$ and $\Sigma = Th \bar{\mathfrak{U}} \cup T'$. Obviously any model of Σ has an isomorph $\mathfrak{U}^{(1)}$ for which (6) holds. To show that Σ has a model, let $\phi_1(c), \dots, \phi_n(c)$ (c not in ϕ_i) belong to T' , and write $\psi = \phi_1 \wedge \phi_2 \wedge \dots \wedge \phi_n$. Then clearly, $\exists v_0 \psi(v_0)$ belongs to $Th \bar{\mathfrak{U}}$. Hence, for some \bar{a} , $\bar{\mathfrak{U}} \vdash \psi[\bar{a}]$ and $(\bar{\mathfrak{U}}, \bar{a})$ is a model of $Th \bar{\mathfrak{U}} \cup \{\phi(c), \dots, \phi_n(c)\}$. Hence, by compactness, Σ has a model, and the proof is complete.

Incidentally, the little argument just made involving the existential quantifier is used many times in model theory.

We stated 4.4 for arbitrary $\bar{S}(\bar{T})$, but our primary concern is when $S(T)$ is countable, as we now assume. 4.4 still does not give a saturated model of T of power \aleph_0 , and indeed there may not be one. For an example, consider the theory T_1 of the field $\mathfrak{U}_1 = (\text{Reals}, +, \cdot)$. In \mathfrak{U}_1 , the usual ordering $<$ is elementarily definable ($x \leq y \leftrightarrow \exists z (z^2 = y - x)$). Clearly, for each rational q , we can introduce a formula $\phi_q(v_0)$ saying that ' $q < v_0$ '. For each real x , let $Q_x = \{q \in \mathbb{Q} : \mathfrak{U}_0 \vdash \phi_q[x]\}$. Then clearly Q_x determines x . Hence, if $x_1 \neq x_2$, then $Th(\mathfrak{U}_1, x_1) \neq Th(\mathfrak{U}_1, x_2)$. Thus T_1 has 2^{\aleph_0} types of elements (even without using parameters). Any countable model of T_1 can realize only countably many types, so there is no denumerable saturated model of T_1 . (Using the same arguments, one can show there are

2^{\aleph_0} non-isomorphic denumerable models of T_1 .)

In order to specify just when T has a denumerable saturated model, we need an obvious generalization of the notion of type of element over T . If \mathfrak{U} is a model of T , and $a_0, \dots, a_{n-1} \in A$ (and c_0, \dots, c_{n-1} are new constants), then $Th(\mathfrak{U}, a_0, \dots, a_{n-1})$ is called a **type of n -tuple of elements**, or simply **n -type, over T** .

THEOREM 4.5. *A countable complete theory T , having infinite models, has a denumerable saturated model if and only if for each n , T has countably many n -types.*

Proof. The necessity is just as in the example of the real field above. For sufficiency, one simply repeats the proof of 4.4. At the point where the number of pairs (x, T') must be shown to be small enough, the old hypothesis $\bar{S} + \aleph_0 < \kappa$ is replaced by our new assumption about the number of n -types.

The field of complex numbers is saturated and so is the denumerable algebraically closed field of characteristic zero and transcendence degree \aleph_0 . More generally, if T is categorical in a non-denumerable power then every non-denumerable model of T can be shown without using the G.C.H. to be saturated (this is part of the proof of Morley's Theorem 3.4); and also, T has a denumerable saturated model. The complete theory T_1 of the real field was shown above to have no denumerable saturated model. Of course, using G.C.H., T_1 has a saturated model in every regular uncountable power \aleph_α . What is special here, is that these saturated models of T_1 can be shown to coincide with the fields already studied (by Erdős, Gillman, and Hendrickson) under the name ' η_α -real closed fields'. If a theory T' is categorical in the power \aleph_0 , then the condition in 4.5 obviously holds, so the unique denumerable model of T' is saturated. (An example is the theory of densely ordered sets without extreme points!)

Some of the remarks just made require definitely nontrivial proofs. By 4.4 and 4.5 there are a great many saturated models. However, to prove that a *given* model \mathfrak{U} is saturated is usually not easy, because (in view of the definition of 'saturated') it is likely to require much information about what an elementary formula can say about an element of \mathfrak{U} , or a pair of elements, etc.

5. Ultraproducts. As is familiar, a family D of subsets of a given set I is called an **ultrafilter** over I if (i) $0 \notin D$, (ii) if $X \in D$ and $X \subseteq Y$ then $Y \in D$, (iii) if $X, Y \in D$ then $X \cap Y \in D$, and (iv) if $X \subseteq I$ then $X \in D$ or $I - X \in D$. It is easy to show, using Zorn's lemma, that

- (1) if F is a family of subsets of I having the finite intersection property (no finite intersection of members of F is empty), then for some ultrafilter D , $F \subseteq D$.

For each $i \in I$, $\{X \subseteq I : i \in X\}$ is clearly an ultrafilter; ultrafilters of this special form are called **principal**. Obviously, D is non-principal if and only if D contains $I - X$

for each finite set X . If I is cofinite the set of all such 'infinite' sets obviously has the finite intersection property, so, by (1), there is at least one non-principal ultrafilter over I .

An ultrafilter D can also (and suggestively) be viewed as a two-valued finitely additive measure μ . (For each $X \subseteq I$, simply put $\mu X = 1$ or 0 according as $X \in D$ or not.)

In this spirit we say that a property $P(i)$ holds for D -almost all i to mean that $\{i \in I: P(i)\} \in D$.

Now suppose that to each $i \in I$ is correlated an S -structure \mathfrak{A}_i , and that D is an ultrafilter over I . Let C be the Cartesian product $\Pi(A_i: i \in I)$ of the sets A_i . If $f, g \in C$, we write $f \sim_D g$ if $f(i) = g(i)$ for D -almost all i . \sim_D is easily seen to be an equivalence relation on C . Let f/D denote the equivalence class of f and put $B = \{f/D: f \in C\}$. We now form an S -structure \mathfrak{B} having universe B and called the *ultraproduct* $\Pi_D(\mathfrak{A}_i: i \in I)$. If $P \in S$ and P is n -ary, then $(*)P^{f_0/D} \dots f_{n-1/D}$ is to hold if and only if $P^{f_0(i) \dots f_{n-1}(i)}$ holds for D -almost all i .

Observe that the ultraproduct operation has been defined purely algebraically — with no mention of logic. The central and remarkable fact about ultraproducts is that, nevertheless, there is an immediate and strong connection between ultraproducts and logic. In fact we can prove that the relationship $(*)$ extends from atomic formulas to all formulas of $L(S)$:

THEOREM 5.1. *Basic ultraproduct theorem (Łoś). Let D , \mathfrak{A}_i , \mathfrak{B} and C be as above. Then for any S -formula with distinct free variables u_0, \dots, u_{n-1} and any $f_0, \dots, f_{n-1} \in C$, $\mathfrak{B} \vdash \phi \left(\begin{smallmatrix} u_0 & \dots & u_{n-1} \\ f_0/D & \dots & f_{n-1}/D \end{smallmatrix} \right)$ if and only if $\mathfrak{A}_i \vdash \phi \left(\begin{smallmatrix} u_0 & \dots & u_{n-1} \\ f_0(i) & \dots & f_{n-1}(i) \end{smallmatrix} \right)$ for D -almost all i .*

It is just an exercise to establish 5.1 by 'induction on the formula ϕ .' Notice that no theorem of logic (beyond the basic definitions in §1) is required in the proof.

The definition of ultraproduct and Theorem 5.1 have been given assuming S contains only relation symbols. However, by 5.1, if P is, e.g., ternary, and each $P^{f_0(i) \dots f_{n-1}(i)}$ is 'operational' (i.e., $\mathfrak{A}_i \vdash (\forall u, v)(\exists! w)Puvw$) then $P^{f_0/D \dots f_{n-1}/D}$ is also operational. Hence the definition of ultraproduct can be modified in an obvious way to allow operation symbols. Thus, as usual, there is no loss of generality in assuming S has only relation symbols.

Using 5.1, we shall now give a second proof (due to D. Scott and A. Tarski) of the compactness theorem 2.1. Suppose that every finite subset s of a given set Σ of S -sentences has a model, say, \mathfrak{A}_s . Let I be the set of all finite subsets of Σ . For each $\sigma \in \Sigma$, let $W_\sigma = \{s \in I: \sigma \in s\}$. Obviously, the set of all such W_σ has the finite intersection property, so by (1) it can be extended to an ultrafilter D over I . Then $\mathfrak{B} = \Pi_D(\mathfrak{A}_s: s \in I)$ is the desired model of Σ . Indeed, if $\sigma \in \Sigma$, then each \mathfrak{A}_s such

that $\sigma \in s$ was taken to be a model of σ . Therefore $\mathfrak{U}_s \vdash \sigma$ for all $s \in W_\sigma$ and so for D -almost all s . Hence, by 5.1, $\mathfrak{B} \vdash \sigma$.

Incidentally, the inquisitive reader who tries to compare this beautiful proof of 2.1 with that given in §2 will find that the two proofs have more in common than first appearances might indicate. For example, looking at 5.1, suppose we expand S to S' by introducing a new constant c_f for each $f \in C$. Let Γ be the set of all S' -sentences true in all the structures $(\mathfrak{U}_i, f(i))_{f \in C}$, for $i \in I$. Then clearly Γ has the property that was possessed by the set Δ_S in the original proof of 2.1, namely: For any 1-formula ϕ over S' , there exists $d \in S'$ such that $\exists v_0 \phi \rightarrow \phi(d) \in \Gamma$.

There are a number of basic facts about ultraproducts whose proofs (sometimes using 5.1) will easily be found by the reader. If D is principal, say, $D = \{X : j \in X\}$ then $\Pi_D(\mathfrak{U}_i : i \in I) \cong \mathfrak{U}_j$. When all \mathfrak{U}_i are the same structure \mathfrak{U} , the ultraproduct is written \mathfrak{U}_D^I and called an **ultrapower**. For $a \in A$, let $h(a)$ be the constant function on I with value a , taken mod D . Then h is an elementary embedding of \mathfrak{U} into \mathfrak{U}_D^I . We often 'identify' \mathfrak{U} and its image under h , so that \mathfrak{U} becomes an elementary substructure of \mathfrak{U}_D^I . If \mathfrak{U} is finite, then $\mathfrak{U}_D^I \cong \mathfrak{U}$ (by 5.1). It is known that, D being non-principal and \mathfrak{U} infinite, \mathfrak{U}_D^I is nearly always a **proper** extension of \mathfrak{U} . For example, consider the case $I = \omega$. Choose any one-to-one function f on I into A . Then for each $a \in A$, $\{i \in \omega : f(i) = a\}$ has at most one element, so $f/D \neq h(a)$. Hence f/D is not in (the canonical image of) A .

This last remark has an application to model theory, of which we will only give an example. Suppose \mathfrak{U} is a structure of power 2^{\aleph_0} . Is there a proper elementary extension of \mathfrak{U} of the same power? If $\bar{S} \leq 2^{\aleph_0}$, then (by considering $Th(\mathfrak{U}, a)_{a \in A} \cup \{c \approx c_a : a \in A\}$) we obtain at once from 3.1 and 3.2 an affirmative answer. But if $\bar{S} > 2^{\aleph_0}$ that argument fails. However, ultraproducts give an affirmative answer. Take a non-principal D over ω . We just saw that \mathfrak{U}_D^I is a proper elementary extension of \mathfrak{U} . Its power is at most that of the Cartesian power A^ω , which is $(2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0}$, as desired. (However, if A has certain cardinal numbers, there is no such extension. For details of the known results on this subject due to Rabin, Keisler, and others, see [3].)

It follows at once from 5.1 that \mathcal{EC} or even \mathcal{EC}_Δ properties \mathcal{K} are preserved by ultraproducts (i.e., if all \mathfrak{U}_i belong to \mathcal{K} so does $\Pi_D(\mathfrak{U}_i : i \in I)$). However, even more can be said, because of the simple but important fact that, in defining the ultraproduct \mathfrak{B} the universe B depends only on the universes A_i and, moreover, for each $P \in S$, $P^{\mathfrak{B}}$ depends only on the $P^{\mathfrak{U}_i}$ (and not on $Q^{\mathfrak{U}_i}$ for other $Q \in S$). Indeed, we have at once: *\mathcal{PC}_Δ properties \mathcal{K} are preserved by ultraproducts.* (If $\mathcal{K} = \mathcal{K}' \upharpoonright S$, expand each \mathfrak{U}_i to $\mathfrak{U}'_i \in \mathcal{K}'$. Then by 5.1, $\Pi_D(\mathfrak{U}'_i : i \in I) \in \mathcal{K}'$ and $\Pi_D(\mathfrak{U}_i : i \in I) = \Pi_D(\mathfrak{U}'_i : i \in I) \upharpoonright S$). For example, an ultraproduct of non-well-ordered structures $(A_i, <_i)$ is not well-ordered.

Ultraproducts are a whole subject in themselves. For example, even the list of 'basic facts' begun above can be continued at some length — before some harder

results are encountered. It is clear from what has already been said that ultraproducts play a role not only in model theory, but also in algebra, and even in analysis (where, after all, the 'almost everywhere' approach was invented). Moreover, a plain set A is of course a structure (with S empty), and when we take ultraproducts $\Pi_D(A_i; i \in I)$ of sets we encounter the extremely interesting and difficult problem in set theory: what can be said about the cardinal number \bar{B} of the ultraproduct? (\bar{B} clearly depends only on the cardinal numbers of the A_i and on D .) Thus ultraproducts also are part of the subject matter of set theory (and certainly ultrafilters are).

Here we can only make this brief introduction to ultraproducts, and recommend the books [2] and [3] for further reading and references. We close by stating without proof two important results concerning ultraproducts in model theory.

A deeper analysis shows that by choosing special kinds of ultrafilters one can make the ultraproducts have various special properties. A notable example is the following theorem of Keisler (cf. [3]).

THEOREM 5.2. (G.C.H) *For any I of infinite power κ , there exists an ultrafilter D over I such that, for any S with $\bar{S} \leq \kappa$ and any S -structures $(\mathfrak{A}_i; i \in I)$ each of infinite power $\leq \kappa^+$, the ultraproduct $\Pi_D(\mathfrak{A}_i; i \in I)$ is saturated and of power κ^+ .*

Related to 5.2 is the following remarkable result which shows that the notion of elementary equivalence can be characterized in a 'purely mathematical' way (i.e., with no mention of a language).

THEOREM 5.3. *\mathfrak{A} and \mathfrak{A}' are elementarily equivalent if and only if they have isomorphic ultrapowers.*

Indeed, if we choose, $\kappa \geq \bar{S}$, \bar{A} , \bar{A}' , and $\bar{I} = \kappa$, then, by 5.2, there exist D and D' over I such that \mathfrak{A}_D^I and $\mathfrak{A}'_{D'}^I$ are both saturated and of power κ^+ ; hence, by 4.3(a) they are isomorphic.

However, that argument depends on 5.2 which assumes the generalized continuum hypothesis. Ten years after Keisler proved 5.2 (and 5.3 using G.C.H.), S. Shelah (in an article to appear in Israel J. Math.) has very recently succeeded in proving 5.3 outright.

It follows easily from 5.3 and 5.1 that various other notions such as ' \mathcal{K} is an \mathcal{EC}_Δ -class' can also be given purely mathematical definition (cf. [3])—i.e., definitions making no mention of a language. The problem of finding such definitions was initiated and emphasized by Tarski.

6. The ω -completeness theorem. Ultraproducts and saturated models tend to be uncountable and, vaguely speaking, to contain many kinds of elements. We

now study a method of obtaining models which are denumerable and omit certain kinds of elements. Throughout this section S is assumed to be countable.

Suppose S contains a unary relation symbol N , and distinct constants $c_0, c_1, \dots, c_n, \dots$ (as well as other symbols). An S -structure \mathfrak{A} is called an ω -model (with respect to N, c_0, c_1, \dots) if $N^{\mathfrak{A}} = \{c_0^{\mathfrak{A}}, \dots, c_n^{\mathfrak{A}}, \dots\}$. For example, let $\mathfrak{A}_0 = (\text{Reals}, +, \cdot, \omega, 0, 1, \dots, n, \dots)$, where $N^{\mathfrak{A}_0} = \omega$ and $c_n^{\mathfrak{A}_0} = n$. If $\mathfrak{B} \equiv \mathfrak{A}_0$ then the 'natural numbers' $N^{\mathfrak{B}}$ of \mathfrak{B} may be 'non-standard' (meaning in this case that \mathfrak{B} is non-Archimedean), in which case \mathfrak{B} is not an ω -model.

As we have seen in examples in §2, in general no set of elementary sentences can "say that $N = \{c_0, c_1, \dots\}$ ", so the notion ' ω -model' is not elementary. Nevertheless, as we will see, under certain conditions we can be sure that T has an ω -model.

Let T be an S -theory such that $T \vdash Nc_n$ for each n . T is called ω -complete if

- (1) for any S -sentence $\forall v_0 \phi$, if for all n , $T \vdash \phi(c_n)$, then $T \vdash \forall v_0 (Nv_0 \rightarrow \phi(v_0))$.

For example, it is obvious that

- (2) the theory of an ω -model \mathfrak{A} or of any class of ω -models is ω -complete.

The basic result called the ω -completeness theorem states that *if T has a model and is ω -complete, then T has an ω -model*.

As it happens, one can by almost exactly the same proof establish a more general statement, having broader applications. (The original ω -completeness theorem is due to Henkin and Orey; the general statement just evolved.) To state the broader form we must first adopt a broader meaning of ' ω -model' and ' ω -complete.' Instead of N and c_0, c_1, \dots , we now are given a fixed list $\theta_0, \theta_1, \dots, \theta_n, \dots$ of 1-formulas of $L(S)$. \mathfrak{A} is called an ω -model (with respect to the list $\theta_0, \theta_1, \dots$) if $A = \bigcup (\theta_n^{\mathfrak{A}} : n \in \omega)$. T is called ω -complete if whenever $T \vdash \forall v_0 (\theta_n(v_0) \rightarrow \phi(v_0))$ for all n , then $T \vdash \forall v_0 \phi(v_0)$. The wording of the ω -completeness theorem is unaltered. To interpret the old ideas in the new, let θ_n be the formula $v_0 \approx c_n \vee \sim Nv_0$. Then clearly the two notions of ω -model coincide, as do the two notions of ω -complete. (In the new notions we have replaced each individual c_n by a set, and then dropped the subuniverse N , which is still covered by the device used just above.)

It is worthwhile noting that in contrapositive form the definition of ω -complete reads:

- (3) T is ω -complete if and only if for any 1-formula ϕ , if $T + \exists v_0 \phi(v_0)$ has a model, then for some n , $T + \exists v_0 (\theta_n(v_0) \wedge \phi(v_0))$ has a model.

In particular, taking $\phi(v_0)$ to be $v_0 \approx d$, we obtain:

- (4) if T is ω -complete and $d \in S$, then for some n , $T + \theta_n(d)$ has a model.

A rather simple but important fact about ω -completeness is in

THEOREM 6.1. *Let σ be a sentence involving the symbols in $S(T)$ and perhaps new individual constants. If T is ω -complete, so is $T + \sigma$.*

Proof. First assume σ has no new constants. Suppose, that for all n , $T + \sigma \vdash \forall v_0(\theta_n(v_0) \rightarrow \phi(v_0))$. Then, clearly, for all n ,

$$T \vdash \forall v_0(\theta_n(v_0) \rightarrow (\sigma \rightarrow \phi(v_0))).$$

Since T is ω -complete, we infer that $T \vdash \forall v_0(\sigma \rightarrow \phi(v_0))$, that is,

$$T + \sigma \vdash \forall v_0 \phi(v_0),$$

as desired.

It is now clearly enough to deal with the case (denoted by $T + c$) where one new constant c and no σ is added. Suppose that for all n , $T + c \vdash \forall v_0(\theta_n(v_0) \rightarrow \phi(v_0))$. We can suppose $\phi(v_0)$ is $\psi\left(\begin{smallmatrix} v_0 & v_1 \\ v_0 & c \end{smallmatrix}\right)$, where c is not in ψ . Then clearly, for all n ,

$$T \vdash \forall v_0(\theta_n(v_0) \rightarrow \forall v_1 \psi(v_0, v_1)).$$

Hence, by the ω -completeness of T , $T \vdash \forall v_0 \forall v_1 \psi(v_0, v_1)$, whence

$$T + c \vdash \forall v_0 \psi\left(\begin{smallmatrix} v_0 & v_1 \\ v_0 & c \end{smallmatrix}\right); \text{ as desired.}$$

Recall that, given $S = S(T)$, we formed in (2) of §2 the set Δ_S consisting of certain sentences $\sigma_n = \exists u_n \phi_n \rightarrow \phi_n\left(\begin{smallmatrix} u_n \\ d_n \end{smallmatrix}\right)$, when d_0, \dots, d_n, \dots are new constants (and now $\kappa = \omega$). By 6.1, each theory $T_m = T + \{\sigma_0, \dots, \sigma_{m-1}\}$ is ω -complete, assuming that T is. An infinite extension usually does not preserve ω -completeness, but the very special extension $T' = T \cup \Delta_S$ does:

LEMMA 6.2. *If T is ω -complete so is $T + \Delta_S$.*

Proof. The point is that for each m , T' is a 'conservative' extension of T_m , i.e., if σ is an $S(T_m)$ -sentence and $T' \vdash \sigma$ then $T_m \vdash \sigma$. Indeed, if $T' \vdash \sigma$ then by compactness, $T_{m+k} \vdash \sigma$ for some k . But any model \mathfrak{A} of T_m can clearly be expanded to a model of T_{m+k} (cf. (1) of §2), so $\mathfrak{A} \vdash \sigma$; thus $T_m \vdash \sigma$, as claimed. Now suppose, for all n , $T' \vdash \forall v_0(\theta_n(v_0) \rightarrow \phi(v_0))$. For some m , ϕ is an $S(T_m)$ -formula, and T' conservatively extends T_m , so $T_m \vdash \forall v_0(\theta_n(v_0) \rightarrow \phi(v_0))$, for all n . Hence, $T' \vdash \forall v_0 \phi(v_0)$, as desired (since T_m is ω -complete and $T_m \subseteq T'$).

Now we are ready for

THEOREM 6.3. (ω -Completeness Theorem). *If T has a model and is ω -complete then T has a (countable) ω -model.*

Proof. The notation above is continued. We define $k_0, k_1, \dots, k_n, \dots$ inductively in such a way that for each n ,

$$(5) \quad \Sigma_n = T' + \{\theta_{k_0}(d_0), \dots, \theta_{k_{n-1}}(d_{n-1})\} \text{ has a model.}$$

It is clear (by (3) of §2) that $\Sigma_0 = T'$ has a model. Suppose we have (5) for a given n . By 6.1 and 6.2, Σ_n is ω -complete. Hence, by (4), we can find k_n such that $\Sigma_n + \{\theta_{k_n}(d_n)\}$ has a model, and the inductive step is complete.

By the compactness theorem, $\Sigma = \bigcup_n \Sigma_n$ has a model and, indeed, by 2.3 (or 3.1), Σ has a model \mathfrak{A} in which $A = \{d_n^{\mathfrak{A}} : n \in \omega\}$. Since for each n , $\mathfrak{A} \models \theta_{k_n}(d_n)$, \mathfrak{A} is an ω -model.

The model we obtained was in fact countable, but we could always obtain a countable model in 6.3 from an arbitrary one by applying the downward Löwenheim-Skolem Theorem.

There are two ways in which 6.3 can still be strengthened. First, the formulas $\theta_n(v_0)$ can be replaced by a list $\theta_n(v_0, \dots, v_k)$ of $(k+1)$ -formulas, for an arbitrary fixed k . All definitions and propositions are altered in the obvious way and the proofs need only trivial changes. The second improvement allows more than one list of θ_n 's, and in fact countably many lists, so we are given $\theta_n^m(v_0, \dots, v_{k(m)})$ ($m, n = 0, 1, \dots$). \mathfrak{A} is called an ω -model (over all these lists) if \mathfrak{A} is an ω -model over each; and T is called ω -complete (over all these lists) if T is ω -complete over each. The ω -completeness theorem in final form reads the same as ever! (6.3 is henceforth given this broader interpretation.) The only change needed in the proof is in the final proof of 6.3. There are now many more things to accomplish (in place of only (5)), but there are still only countably many, so our tasks can be arranged in an ω -list and then accomplished exactly as before.

The condition for \mathfrak{A} to be an ω -model with respect to one list of k -formulas $\theta_0, \dots, \theta_n, \dots$ is just that \mathfrak{A} have no k -tuple of elements satisfying $\Phi = \{\sim \theta_n : n \in \omega\}$; or, as we say, that \mathfrak{A} has no k -tuple of elements of the kind Φ ; or simply that \mathfrak{A} omits Φ . Therefore 6.3 is sometimes called the 'omitting kinds theorem.' (Actually the word 'type' is used instead of 'kind,' but in §4 we have reserved that for the case when (the theory axiomatized by) Φ is complete.)

6.3 and 6.1 together imply a converse of (2):

$$(6) \quad \text{If } T \text{ is } \omega\text{-complete and } \mathcal{K} \text{ is the class of all } \omega\text{-models of } T, \text{ then } T = Th \mathcal{K}.$$

Indeed, if $\sigma \notin T$, then $T + \sim \sigma$ is ω -complete, by 6.1, so has an ω -model by 6.3; hence $\sigma \in Th \mathcal{K}$.

Speaking informally now, if we are given a set Σ of sentences, we can consider the smallest set $\Sigma' \supseteq \Sigma$ closed under the axioms and rules of inference of first order logic (see the discussion just after 2.3) and also under the (infinitary) " ω -rule": if $\phi(c_n) \in \Sigma'$ for each n , then $\forall v_0 (Nv_0 \rightarrow \phi(v_0)) \in \Sigma'$. (Here we are returning for simplicity to the simplest notions of ω -model and ω -complete, that is, over N, c_0, c_1, \dots .) Then (6) tells us that

- (7) a sentence σ belongs to Σ' if (and only if) σ is true in all ω -models of Σ .

It is in the form (7) that the ω -completeness theorem gets its name — as an analogue of Gödel's completeness theorem. The general ω -completeness theorem is also, in a sense, equivalent to the completeness theorem for $L_{\omega_1\omega}$ of Karp [5], a very useful completeness result concerning certain (infinitary) axioms and (infinitary) rules of inference for the language $L_{\omega_1\omega}$. (For one statement on a connection between $L_{\omega_1\omega}$ and ω -models, see 7.5.)

There are many interesting applications of the ω -completeness theorem. One of the nicest will be the topic of §7. Another, which we now discuss, deals with denumerable models of complete theories and especially with prime models and \aleph_0 -categoricity. *We assume henceforth in this section that T is a complete S -theory having infinite models.*

We shall first prove a result about omitting types of elements over (a complete theory) T . Let us fix a list of new constants c_0, \dots, c_n, \dots not in $S = S(T)$. As in §4, an n -type (over T) is any complete theory T' over $S_n = S \cup \{c_0, \dots, c_{n-1}\}$. If T' is finitely axiomatizable over T so that, in fact, for some single n -sentence σ , $T' = T + \sigma$, then we say T' is **principal**. Moreover, if ϕ is an n -formula over S , and $T + \phi(c_0, \dots, c_{n-1})$ is complete, then ϕ is called an n -atom (over T). (In fact, ϕ is an atom in the Boolean algebra of n -formulas modulo T -equivalence.) On the other hand, ϕ is called *atomless* (over T) if $T \vdash (\exists v_0, \dots, v_{n-1})\phi$ and for no n -atom ψ do we have $T \vdash \psi \rightarrow \phi$.

THEOREM 6.4. *Any complete theory T has a denumerable model \mathfrak{A} in which every finite sequence of elements satisfies either an atom or an atomless formula. Moreover, given non-principal k_m -types T'_m ($m = 0, 1, \dots$), such an \mathfrak{A} can be found which omits every T'_m .*

Proof. Consider any fixed m , and let ψ_n ($n \in \omega$) be a list of the k_m -formulas such that $\psi_n(c_0, \dots, c_{m-1}) \in T'_m$. Then T is ω -complete with respect to $\sim\psi_0, \dots, \sim\psi_n, \dots$. To see this first note that $T \not\vdash \forall v_0 \psi$ if and only if $T \vdash \exists v_0 \neg\psi$, since T is complete. Assume $k_m = 1$ to simplify notation. Now suppose that for each n , $T \vdash \sim\psi_n(v_0) \rightarrow \phi(v_0)$, but $T \vdash \exists v_0 \sim\phi$. Then clearly $\sim\phi(c_0)$ would be an axiom for T'_m , contrary to the assumption T'_m is non-principal.

Secondly, let k be fixed, and let ϕ_n ($n \in \omega$) be a list of all k -formulas such that ϕ_n

is either atomless or an atom. Again, T is ω -complete with respect to $\phi_0, \dots, \phi_n, \dots$. Indeed, suppose ϕ is a k -formula over S and, for each n , $T \vdash \phi_n \rightarrow \phi$, but $T \vdash (\exists v_0, \dots, v_{k-1}) \sim \phi$. Then for any k -atom ψ , $T \not\vdash \psi \rightarrow \sim \phi$; so $\sim \phi$ is atomless. Hence $T \vdash \sim \phi \rightarrow \phi$, so $T \vdash \phi$, a contradiction.

Now we apply the general ω -completeness theorem to the ' $\omega + \omega$ ' lists above (one for each m , and one for each k), and obtain a denumerable model \mathfrak{A} , ω -complete with respect to every list. Clearly \mathfrak{A} is just as desired.

A model \mathfrak{A} of T is called **prime** if \mathfrak{A} is elementarily embeddable in every model of T . \mathfrak{A} is called **atomic** if, for any n , every n -tuple of elements of A satisfies an n -atom (i.e., for each n , \mathfrak{A} omits every non-principal n -type). The theory T will be called **atomistic** if for each n , there are no atomless n -formulas over T .

THEOREM 6.5. *Let \mathfrak{A} and \mathfrak{B} be models of T .*

- (a) *If \mathfrak{A} is denumerable and atomic then \mathfrak{A} is homogeneous.*
- (b) *If \mathfrak{A} and \mathfrak{B} are denumerable and atomic, then $\mathfrak{A} \cong \mathfrak{B}$.*
- (c) *\mathfrak{A} is prime if and only if \mathfrak{A} is denumerable and atomic.*
- (d) *T has a prime model if and only if T is atomistic.*

Proof. We shall use 6.4 and also some simple Cantor-type arguments (see (1), (2), and (3) of §4), the latter being left to the reader. (a) is trivial using 4.3(b). (b) is easy using Cantor's method. As regards (c), if \mathfrak{A} is denumerable and atomic, and $\mathfrak{A} \equiv \mathfrak{B}$, then it is easy to embed \mathfrak{A} elementarily in \mathfrak{B} , by using Cantor's method; so \mathfrak{A} is prime. Now assume \mathfrak{A} is prime. By the Löwenheim-Skolem theorem, T has a countable model, so clearly \mathfrak{A} is denumerable. If \mathfrak{A} is not atomic, then, for some $a_0, \dots, a_n \in A$, $T' = Th((\mathfrak{A}, a_0, \dots, a_n))$ is non-principal. By (a very special case of) 6.4, T has a model \mathfrak{B} omitting T' . But then clearly \mathfrak{A} cannot be elementarily embedded in \mathfrak{B} , a contradiction. In (d), if T has a prime, and hence atomic, model, it follows trivially that T is atomistic. Finally, suppose T is atomistic. By 6.4, there is a denumerable model \mathfrak{A} of T in which every finite sequence of elements satisfies an atom or an atomless formula. Since the second possibility can never occur, \mathfrak{A} must be atomic, and hence (by (c)) prime.

We need below two simple facts about n -types which are really just familiar results concerning arbitrary Boolean algebras.

- (8) *If there are infinitely many T -inequivalent n -formulas, then there is a non-principal n -type over T .*
- (9) *If there is an atomless n -formula over T , then there are uncountably many (in fact 2^{\aleph_0}) n -types over T .*

Proof. As regards (8), if there is an atomless m -formula ϕ over T then any m -type containing $\phi(c_0, \dots, c_{n-1})$ is clearly non-principal. If there is not then there are obviously infinitely many T -inequivalent n -atoms. Let $\Sigma = \{\sim \phi(c_0, \dots, c_{n-1})$:

ϕ is an m -atom}. Clearly $T \cup \Sigma$ is finitely satisfiable, so has a model $(\mathfrak{A}, a_0, \dots, a_{n-1})$. Then $T' = Th(\mathfrak{A}, a_0, \dots, a_{n-1})$ is non-principal.

Regarding (9), we only remark that given an atomless ϕ we can clearly find atomless formulas ϕ_0 and $\phi_1 = \phi \wedge \sim \phi_0$ such that $T \vdash \phi_0 \rightarrow \phi$; i.e., we can split ϕ in half. Iterating indefinitely, we easily obtain (by another famous argument going back to Cantor) the desired 2^{\aleph_0} different n -types.

A direct consequence of (9) is

(10) *If T has a denumerable saturated model, then T has a prime model.*

Indeed, by 4.5, T has only countably many n -types, for each n . Hence T is atomistic, by (9), and so T has a prime model, by 6.5(d).

THEOREM 6.6. *T is \aleph_0 -categorical if and only if for each n , there are only finitely many T -inequivalent n -formulas (or what is the same, there are no non-principal n -types).*

Proof. The second condition and the parenthetical condition are equivalent by (8) (and its trivial converse). To say there are no non-principal n -types, for any n , is just to say that (*) all denumerable models of T are atomic. (*) implies \aleph_0 -categoricity by 6.5(b). If T is \aleph_0 -categorical then by (a very special case of) 6.4, T clearly cannot have a non-principal n -type. (Or, its only denumerable model clearly satisfies the definition of 'prime', so by 6.5(c), (*) holds.)

From 6.4–6.6 together with what we know about denumerable saturated models follows a curious result:

THEOREM 6.7. *No complete theory T has exactly two nonisomorphic denumerable models.*

Ehrenfeucht gave examples of complete theories T_n having exactly n such models ($n = 3, 4, 5, \dots$)—cf. [17].

Proof. Suppose T has exactly two. Then clearly there are only countably many n -types over T , for each n . Hence T has a denumerable saturated model \mathfrak{B} (by 4.5) and a prime model \mathfrak{A} .

Since T is not categorical in power \aleph_0 , there is (by 6.6) a non-principal n -type T' over T . Of course, T' is realized in \mathfrak{B} , say by (b_0, \dots, b_{n-1}) , but not in \mathfrak{A} , so $\mathfrak{A} \not\cong \mathfrak{B}$. Since T has infinitely many inequivalent n -formulas, so obviously has T' . Hence (again by 6.6), T' is not \aleph_0 -categorical. Hence T' has a denumerable model $(\mathfrak{C}, x_0, \dots, x_{n-1})$ not isomorphic to $(\mathfrak{B}, b_0, \dots, b_{n-1})$. But \mathfrak{B} is homogeneous (by 4.3(c)), so clearly $\mathfrak{C} \not\cong \mathfrak{B}$. Of course, $\mathfrak{C} \not\cong \mathfrak{A}$, as \mathfrak{C} realizes T' . This is a contradiction.

The theory T_0 of the complex number field has a denumerable saturated model

and hence (by (10)) a prime model. (In fact the field of complex algebraic numbers is the prime model of T_0 .) However, the theory T_1 of the real field has no denumerable saturated model (cf. §4), but does have a prime model, the field of algebraic real numbers (although this will not be proved here). If \mathfrak{A} is any structure which has a definable well-ordering (e.g., $(\omega, +, \cdot)$), then obviously $T = Th \mathfrak{A}$ is atomistic and hence has a prime model. This is a very special case, as here the atoms (say, 1-atoms) ϕ satisfy the additional condition $T \vdash \exists! v_0 \phi(v_0)$ and correspond to 'definable elements'. Actually, it takes a little effort to show there is any theory T having no prime model; though it may be that in some sense 'most' theories T do not have prime models.

6.4–6.7 is the work of Ryll-Nardzewski, Ehrenfeucht, Engeler, Svenonius, and the author (cf. [17]).

7. Two-cardinal theorems. The Löwenheim-Skolem theorems (§3) are about the cardinal numbers of the models \mathfrak{A} of an arbitrary set Σ of sentences. Now, in fact, if θ is any fixed 1-formula, then each \mathfrak{A} determines a pair $(\bar{A}, \bar{\theta}^{\mathfrak{A}})$ of cardinals. For example, if the \mathfrak{A} 's are groups, $\theta^{\mathfrak{A}}$ might be (for each \mathfrak{A}) the center of \mathfrak{A} . What can be said about the pairs of cardinals $(\bar{A}, \bar{\theta}^{\mathfrak{A}})$ achieved by models \mathfrak{A} of Σ ? If Σ has a model where the pair is (κ, λ) , for what (κ', λ') can we always be sure that Σ has a model whose pair is (κ', λ') ? Answers or partial answers to these questions may be called Löwenheim-Skolem theorems for two cardinals, or simply, two-cardinal theorems. Incidentally, in most of this work, only infinite $\theta^{\mathfrak{A}}$ are considered, as most questions about finite $\theta^{\mathfrak{A}}$ reduce trivially to questions about one cardinal or about infinite $\theta^{\mathfrak{A}}$.

Again in this section, S is always countable. θ will always be a 1-formula (over S). With θ fixed, we agree to say that \mathfrak{A} is of type (κ, λ) to mean that $\bar{A} = \kappa$ and $\bar{\theta}^{\mathfrak{A}} = \lambda$.

The following two-cardinal theorem was established in [11] by using the homogeneous models of §4.

- (1) If a set Σ of S -sentences has a model \mathfrak{A} of type (κ, λ) , where $\aleph_0 \leq \lambda < \kappa$, then Σ has a model \mathfrak{C} of type (\aleph_1, \aleph_0) .

Later Keisler [6] gave a different proof, using the ω -completeness theorem, and greatly strengthened (1), essentially extending it to the language $L_{\omega_1\omega}$ (see 7.6 below). This work will be the main topic of this section. We start right off with the principal theorem, which is still about the elementary language, but is in such a strong form that it will immediately imply the extension of (1) to $L_{\omega_1\omega}$.

THEOREM 7.1. Suppose $\aleph_0 \leq \bar{\theta}^{\mathfrak{A}} < \bar{A}$. Then there exist \mathfrak{B} and \mathfrak{C} such that $\mathfrak{B} < \mathfrak{A}$, $\mathfrak{C} > \mathfrak{B}$, $\theta^{\mathfrak{B}} = \theta^{\mathfrak{C}}$, $\bar{B} = \aleph_0$, and $\bar{C} = \aleph_1$.

Proof. There is clearly no loss of generality in assuming that $S = S(\mathfrak{A})$ contains a unary relation symbol N such that $\theta^{\mathfrak{A}} = N^{\mathfrak{A}}$. (This is like passing from the group (G, \circ) to the group $(G, \circ, {}^{-1})$.) Let $\kappa = \bar{N}^{\mathfrak{A}}$. By 3.2, \mathfrak{A} has an elementary

substructure \mathfrak{U}' of power κ^+ with the same N (i.e., $N^{\mathfrak{U}} = N^{\mathfrak{U}'}$). Hence we may as well just assume to begin with that $\bar{A} = \kappa^+$.

Let $<$ be a well-ordering of A of order type κ^+ , and let \triangleleft be a corresponding new relation symbol. We introduce the abbreviation $Qu\phi$ (read 'there exist arbitrarily large u such that ϕ ') for the formula $\forall z \exists u (z \triangleleft u \wedge \phi)$ (where z does not occur in ϕ). The cardinal κ^+ is regular (see end §1). Hence the structure $(\mathfrak{U}, <)$ is clearly a model of every sentence of the form:

$$(2) \quad (\forall w_0 \cdots w_{n-1}) [Qu \exists v (Nv \wedge \psi(u, v)) \rightarrow Ev (Nv \wedge Qu\psi(u, v))].$$

(Here ψ is any $S \cup \{<\}$ -formula, perhaps involving the w_i 's as parameters.)

By the downward Löwenheim-Skolem theorem, $(\mathfrak{U}, <)$ has a denumerable elementary substructure $\mathfrak{B}' = (\mathfrak{B}, <)$.

LEMMA 7.2. *Let Σ_0 consist of all the sentences (2), plus a sentence saying that $<$ is a simple ordering of the universe with no last element, plus sentences saying N has at least n elements ($n = 1, 2, 3, \dots$). Then:*

- (a) *Every denumerable model \mathfrak{B}' of Σ_0 has a proper elementary extension with the same N .*

Obviously the \mathfrak{B}' constructed above is a model of Σ_0 . For later reference, we carry out the next steps (Lemma 7.2(a), (b)) in the proof of 7.1 for an arbitrary denumerable model \mathfrak{B}' of Σ_0 .

7.2(a) is the main step in the proof of 7.1. To prove it, first select distinct new constants c_n for $n \in \omega$ and d_b for $b \in B'$. Clearly $N^{\mathfrak{B}'}$ is denumerable, so we write $N^{\mathfrak{B}'} = \{a_0, \dots, a_n, \dots\}$. Let $\mathfrak{B}^* = (\mathfrak{B}', a_n, b)_{n \in \omega, b \in B'}$ (where c_n denotes a_n and d_b denotes b). Select another new constant e , and let $T = Th(\mathfrak{B}^*) + \{c_b \triangleleft e : b \in B'\}$. It is easy to see (without using (2)) that if ψ is any 1-formula over $S(\mathfrak{B}^*)$ then:

$$(3) \quad T + \psi(e) \text{ has a model if and only if } \mathfrak{B}^* \vdash Qu\psi(u).$$

Indeed, if the right side of (3) fails, then for some $b \in B'$, $\mathfrak{B}^* \vdash (\forall v_0 \triangleright d_b) \sim \psi(v_0)$, so $T \vdash \sim \psi(e)$. On the other hand, if $\mathfrak{B}^* \vdash Qu\psi(u)$, then clearly every finite subset of $T + \psi(e)$ has a model (of the form (\mathfrak{B}^*, b')); so $T + \psi(e)$ has a model.

Clearly T has a model (e.g., by (3) with $u \approx u$ for ψ).

We now consider the notions ω -model and ω -complete, with respect to N and c_0, \dots, c_n, \dots (and thus in the simplest form, at the beginning of §6). We will show that T is ω -complete. Suppose ϕ is any 2-formula over $S(T)$ not involving e and

$$(4) \quad T + \exists v (Nv \wedge \phi(e, v)) \text{ has a model.}$$

What we need to show is that for some m , $T + \phi(e, c_m)$ has a model (cf. (1) of §6 in contrapositive form). Now, by (3) and (4),

$$\mathfrak{B}^* \vdash Qu \exists v (Nv \wedge \phi(u, v)).$$

Since \mathfrak{B}' is a model of the sentences (2), we can infer:

$$\mathfrak{B}^* \vdash \exists v(Nv \wedge Qu\phi(u, v)).$$

(ϕ may involve c_i 's and d_b 's, but parameters were allowed in (2).) Hence, for some m , $\mathfrak{B}^* \vdash Qu\phi(u, c_m)$. Therefore, by (3), $T + \phi(e, c_m)$ has a model, as desired.

We have shown that T is ω -complete and has a model so, by the ω -completeness theorem, T has an ω -model \mathfrak{C} . \mathfrak{B}^* is elementarily embeddable in \mathfrak{C} , so we can clearly assume that outright $\mathfrak{B}^* < \mathfrak{C}$. Thus \mathfrak{C} can be written as $(\mathfrak{B}'', a_n, b, x)_{n \in \omega, b \in B'}$, where $\mathfrak{B}' < \mathfrak{B}''$. By the axioms of T , $x \notin B'$, so \mathfrak{B}'' is a proper extension of \mathfrak{B}' . Since \mathfrak{C} is an ω -model, $N^{\mathfrak{B}''} = \{a_n : n \in \omega\} = N^{\mathfrak{B}'}$. Thus 7.2(a) is proved.

LEMMA 7.2(b). *Any denumerable model \mathfrak{B}' of Σ_0 has an elementary extension \mathfrak{C}' of power \aleph_1 with the same N .*

To prove 7.2(b), define denumerable models \mathfrak{B}'_α of Σ_0 for $\alpha < \omega_1$ by recursion as follows: $\mathfrak{B}'_0 = \mathfrak{B}'$. $\mathfrak{B}'_{\alpha+1}$ is a denumerable proper elementary extension of \mathfrak{B}'_α with the same N , as guaranteed by 7.2(a). $\mathfrak{B}'_\alpha = \bigcup (\mathfrak{B}'_\beta : \beta < \alpha)$ if α is a limit ordinal; \mathfrak{B}'_α is a model of Σ_0 by the union theorem (4.1). Then clearly $\mathfrak{C}' = \bigcup (\mathfrak{B}'_\alpha : \alpha < \omega_1)$ is as desired in 7.2(b).

Just before 7.2(a) we constructed a particular $\mathfrak{B}' = (\mathfrak{B}, <)$. Taking $\mathfrak{C}' = (\mathfrak{C}, <')$ as in 7.2(b) for this \mathfrak{B}' , it is clear that \mathfrak{B} and \mathfrak{C} are as demanded in 7.1. Thus 7.1 is proved.

From the proof of 7.1 we can also infer a new compactness theorem and a new completeness theorem. Again it is convenient to assume outright that S contains a special symbol N (instead of dealing with a special formula ϕ) and is just as general. (The type (κ, λ) of \mathfrak{A} refers now to $\bar{A}, \bar{N}^{\mathfrak{A}}$.) Let \triangleleft be a symbol not in S , and let the set Σ_0 be as in 7.2. (Note that Σ_0 depends only on S .)

LEMMA 7.3. *A set Σ of S -sentences has a model of type (\aleph_1, \aleph_0) if and only if $\Sigma \cup \Sigma_0$ has a model.*

Proof. Any model \mathfrak{A} of type (\aleph_1, \aleph_0) can trivially be expanded to a model of Σ_0 (as we saw at the beginning of the proof of 7.1). Hence from left to right is obvious. Now suppose $\Sigma \cup \Sigma_0$ has a model and hence a denumerable model \mathfrak{B}' . By 7.2(b), \mathfrak{B}' has an elementary extension $\mathfrak{A}' = (\mathfrak{A}, <)$ with the same N and of power \aleph_1 . Then \mathfrak{A} is a model of Σ of type (\aleph_1, \aleph_0) .

THEOREM 7.4. (a) (*Compactness*): *If every finite subset of a set Σ of S -sentences has a model \mathfrak{A} of type (\aleph_1, \aleph_0) , then so has Σ .*

(b) (*Completeness*): *An S -sentence σ is true in all models of type (\aleph_1, \aleph_0) if and only if $\Sigma_0 \vdash \sigma$ (and hence if and only if σ is formally derivable from Σ_0 by using the axioms and rules of inference of elementary logic).*

Proof. (a) is immediate from 7.3 and the ordinary compactness Theorem 3.1. (b) is really a special case of 7.3. Indeed, $\Sigma_0 \vdash \sigma$ if and only if $\{\sim \sigma\} \cup \Sigma_0$ has no model

and hence (by 7.3) if and only if σ is true in all models of type (\aleph_1, \aleph_0) . Of course, the parenthetical addition in (b) depends on the completeness theorem, discussed after 2.3.

The two-cardinal theorem (1) is an immediate consequence of 7.1. ((1) and also 7.4 and 7.3 (for a different Σ_0) were originally obtained (by the author and Fuhrken) by a different method, involving homogeneous models.) Let us form a new language $L^*(S)$ by adding to $L(S)$ a new quantifier symbol Q^* and interpreting formulas of the form $Q^* u \phi(u)$ to mean "there are uncountably many u such that $\phi(u)$." By means of a method of translation introduced by Fuhrken, one can infer from 7.4(a) that the ordinary compactness theorem 2.1, holds for $L^*(S)$ as long as S is countable. (1) and 7.4(b) also have implications for $L^*(S)$. (For all this work, and references, see [2].) However, in the completeness theorem for $L^*(S)$ so obtained (and in 7.4(b) itself), the formal 'derivations' of the $L^*(S)$ -valid sentences involve extraneous symbols (e.g., for the above Σ_0 , the symbol \triangleleft)—a feature not present in most logical systems. Keisler has given for $L^*(S)$ a completeness theorem for some simple and elegant axioms and rules of inference involving no extraneous symbols. To do so, he must discard 7.1–7.4 and begin again, working in $L^*(S)$, but some of the arguments are quite close to those for 7.1–7.4 in flavor. (For this interesting work see [7].)

Now we shall see the extra strength of 7.1, over what is in (1) and 7.4, by establishing in 7.6 that (1) holds (for countable Σ) in $L_{\omega_1\omega}$.

First we need the following simple lemma relating $L_{\omega_1\omega}$ to the notion ω -model.

LEMMA 7.5. *For any $L_{\omega_1\omega}(S)$ -sentence σ , there exists a countable set Σ of elementary sentences, containing new symbols $N', d_0, \dots, d_n, \dots$ (and other new symbols as well as those of S) such that for any S -structure \mathfrak{A} :*

\mathfrak{A} is a model of σ if and only if \mathfrak{A} can be expanded to an ω -model of Σ .

Proof. Of course, ' ω -model' is understood with respect to $N', d_0, \dots, d_n, \dots$. Clearly we may assume the only logical symbols in σ are \sim , \exists , and \vee (as well as \approx). Let $N', d_0, \dots, d_n, \dots$ be new symbols. For each subformula ϕ of σ which is a disjunction, introduce a new $(n+1)$ -ary relation symbol P_ϕ , where n is the number of free variables in ϕ . By recursion we correlate with each subformula ϕ of σ a formula ϕ^* . If ϕ is atomic, $\phi^* = \phi$. Also, $(\sim \phi)^* = \sim \phi^*$ and $(\exists u \phi)^* = \exists u \phi^*$. If ϕ is $\vee_n \phi_n$ and its free variables are u_0, \dots, u_{k-1} , then ϕ^* is $\exists z (N'z \wedge P_\phi(z, u_0, \dots, u_{k-1}))$. Σ consists of (i) all the sentences

$$(\forall u_0, \dots, u_{k-1}) (P_\phi(d_m, u_0, \dots, u_{k-1}) \leftrightarrow \phi_m^*)$$

(where $\phi = \vee_n \phi_n$ is a subformula of σ with the free variables u_0, \dots, u_{k-1} , and $m \in \omega$), (ii) all the sentences $d_i \approx d_j$ and $N'd_i$, and (iii) the sentence σ^* . The notation is complicated, but the idea is very simple, and it is easily checked that Σ is as demanded.

THEOREM 7.6. *If a sentence σ of $L_{\omega, \omega}(S)$ has a model \mathfrak{A} of type (κ, λ) where $\aleph_0 \leq \lambda < \kappa$, then σ has a model \mathfrak{C} of type (\aleph_1, \aleph_0) .*

Proof. We can assume that the notion of type (κ, λ) is with respect to a fixed symbol N (i.e., $\lambda = \bar{N}^{\mathfrak{A}}$). Let Σ be as in 7.5. Then \mathfrak{A} can be expanded to an ω -model \mathfrak{A}' (relative to N', d_0, \dots) of Σ . Let θ be $Nv_0 \vee N'v_0$. Obviously $\bar{\theta}^{\mathfrak{A}'} < \bar{A}'$, so we can apply 7.1 to \mathfrak{A}' and θ . We obtain \mathfrak{B}' and \mathfrak{C}' such that $\mathfrak{B}' < \mathfrak{A}'$, $\mathfrak{C}' > \mathfrak{B}'$, $\theta^{\mathfrak{B}'} = \theta^{\mathfrak{C}'}$, $\bar{B}' = \aleph_0$, and $\bar{C}' = \aleph_1$. It is easily verified that \mathfrak{C}' is an ω -model of Σ . Hence, by 7.5, $\mathfrak{C} = \mathfrak{C}' \upharpoonright S$ is a model of σ . Clearly \mathfrak{C} is of type (\aleph_1, \aleph_0) , as desired.

A beautiful application of 7.6, due to Keisler [6], gives a new, short proof of a part of Morley's Theorem 3.7 on theories categorical in power and also strengthens that part of Morley's Theorem. Of course, a class \mathcal{K} of S -structures is called categorical in the power κ if it has, up to isomorphism, exactly one member of power κ .

COROLLARY 7.7. *If a $\mathcal{P}\mathcal{C}_\delta$ -class \mathcal{K} is categorical in power \aleph_1 , then \mathcal{K} is categorical in every uncountable power.*

Proof. We are given that $\mathcal{K} = \mathcal{L} \upharpoonright S$ where $\mathcal{L} = \text{Mod } \Sigma$ and Σ is a set of S_1 -sentences, $S_1 \supseteq S$. (The meaning of ' $\mathcal{P}\mathcal{C}_\delta$ ' (rather than ' $\mathcal{P}\mathcal{C}_\Delta$ ') is that S_1 is assumed countable.) Morley's Theorem 3.8 applied only to $\mathcal{E}\mathcal{C}_\delta$ -classes, so we are extending the ' \aleph_1 -case' of 3.5 to $\mathcal{P}\mathcal{C}_\delta$ -classes (which are much broader).

By (4) of §4, assuming the continuum hypothesis, there is a saturated model \mathfrak{C} of Σ of power \aleph_1 . By (4) of §4, $\mathfrak{C} \upharpoonright S$ is also saturated. Hence (*) there is a saturated member of \mathcal{K} of power \aleph_1 . By using a result in Morley's work, (*) can be obtained without the continuum hypothesis, but the proof we give of (*) (and hence 7.7) will have to depend on the continuum hypothesis.

Let $\kappa > \aleph_1$. There is a member of \mathcal{K} of power κ (by a trivial extension of 3.3 to $\mathcal{P}\mathcal{C}_\delta$ classes). Since \mathcal{K} is categorical in the power \aleph_1 , all members of \mathcal{K} of power κ are elementarily equivalent (by an obvious extension of 3.4 to $\mathcal{P}\mathcal{C}_\delta$). Hence, if all members of \mathcal{K} of power κ are saturated, then they are all isomorphic, by 4.3(a). Now assume \mathcal{K} is not categorical in power κ . Our argument shows that *there is a non-saturated $\mathfrak{A} \in \mathcal{K}$ of power κ* . We plan to infer, using 7.6, that there is a non-saturated member of \mathcal{K} of power \aleph_1 . Since by (*), \mathcal{K} has also a saturated model of power \aleph_1 , this will be a contradiction.

Since $\mathfrak{A} \in \mathcal{K}$, $\mathfrak{A} = \mathfrak{B} \upharpoonright S$ for some $\mathfrak{B} \in \text{Mod } \Sigma$. As \mathfrak{A} is not saturated, there is a subset X of A of power $< \kappa$ and a set Φ of 1-formulas which is finitely satisfiable, but not satisfiable in $(\mathfrak{A}, x)_{x \in X}$. For convenience we can of course take X to be infinite. Write c_x for the symbol denoting x in $(\mathfrak{A}, x)_{x \in X}$. We shall adjoin to \mathfrak{A} countably many new relations which express or code information about Φ .

Each S -formula ϕ is an n -formula for a smallest $n = n(\phi)$, and corresponding to ϕ , we introduce a new relation R_ϕ over X (in \mathfrak{A}) as follows:

$R_\phi(x_1, \dots, x_{n-1})$ if and only if the formula $\phi(v_0, c_{x_1}, \dots, c_{x_{n-1}}) \in \Phi$. Now consider

the structure $\mathfrak{B}^* = (\mathfrak{B}, X, R_\phi)_{\phi \in F}$ (where F is the set of all S -formulas ϕ). Since Φ is finitely satisfiable in $(\mathfrak{U}, x)_{x \in X}$, we have, for any ϕ_0, \dots, ϕ_k (writing $m_i = n(\phi_i) - 1$)

$$(1) \quad \mathfrak{B}^* \vdash \bigwedge_{i \leq k} R_{\phi_i}(u_1^i, \dots, u_{m_i}^i) \rightarrow \exists v_0 \bigwedge_{i \leq k} \phi_i(v_0, u_1^i, \dots, u_{m_i}^i)$$

(where the u_j^i 's are distinct variables).

Since the whole Φ is not finitely satisfiable, we have:

$$(2) \quad \mathfrak{B}^* \vdash \sim \exists v_0 \bigwedge_{\phi \in F} (\forall v_1, \dots, v_{n(\phi)-1}) (R_\phi(v_1, \dots, v_{n(\phi)-1}) \rightarrow \phi(v_0, v_1, \dots, v_{n(\phi)-1})).$$

Since S is countable, F is countable, so the sentence in (2) is in $L_{\omega_1\omega}$. Let σ be the conjunction of (i) all the universalizations of the formulas in (1) (over all possible ϕ_0, \dots, ϕ_k), (ii) the sentence in (2), and (iii) all the sentences in $Th(\mathfrak{B}^*)$. Again, σ is in $L_{\omega_1\omega}$.

By 7.6, there is a model $\mathfrak{B}'^* = (\mathfrak{B}', X', R'_\phi)_{\phi \in F}$ of σ such that $\bar{B}' = \aleph_1$ and $\bar{X}' = \aleph_0$. Let $\mathfrak{U}' = \mathfrak{B}' \upharpoonright S$. Clearly \mathfrak{U}' is a member of \mathcal{K} of power \aleph_1 . Consider $(\mathfrak{U}', x')_{x' \in X'}$, where, say, $d_{x'}$ denotes $x'(x' \in X')$. Let Δ be the set of all $\phi(v_0, d_{x'_1}, \dots, d_{x'_m})$ such that $R'_\phi x'_1 \dots x'_m$ (and $\phi \in F$, $m = n(\phi) - 1$, $x'_i \in X'$). Δ is finitely satisfiable in $(\mathfrak{U}', x')_{x' \in X'}$, because each of the sentences in (1) holds in \mathfrak{B}'^* . On the other hand, since the sentence in (2) holds in \mathfrak{B}'^* , it is clear that Δ is not satisfiable in $(\mathfrak{U}', x')_{x' \in X'}$. Thus (noting that $\bar{X}' < \bar{A}'$) we have shown that \mathfrak{U}' is not saturated, and the proof is complete.

There are more results and some interesting open problems concerning possible further extension of Morley's theorem 3.8 to $\mathcal{P}\mathcal{C}_\delta$ or even $\mathcal{P}\mathcal{C}(L_{\omega_1\omega})$ classes; see [6]. Of course, a $\mathcal{P}\mathcal{C}(L_{\omega_1\omega})$ class is one of the form $(\text{Mod } \sigma') \upharpoonright S$ where σ' is an $L_{\omega_1\omega}$ sentence of type $S' \geq S$. A good example of a class which is (easily seen to be) $\mathcal{P}\mathcal{C}(L_{\omega_1\omega})$ is the class of all free groups (or free anything). Notice that 7.6 immediately implies its own strengthening in which "is a model of σ " is replaced by "belongs to \mathcal{K} ", \mathcal{K} being any $\mathcal{P}\mathcal{C}(L_{\omega_1\omega})$ class.

Let us return now to the simplest two-cardinal theorem (1). In 7.6 we have kept fixed the special role of \aleph_1 in (1) while passing to a language richer than the elementary. What if we keep the elementary language, but change \aleph_1 ? We shall give no proofs but only say something about what is known. By using saturated models plus an ingenious device, Chang proved (cf. [2] or [3]):

THEOREM 7.8 (Chang) (G.C.H.). *If a set Σ of S -sentences has a model of type (\aleph_1, \aleph_0) then Σ has a model of type (κ^+, κ) , provided κ is any regular (infinite) cardinal.*

The exceptional case when κ is not regular remained in total darkness until very recently, when Jensen answered it assuming Gödel's axiom of constructibility. That axiom, usually written ' $V = L$ ', is an axiom much stronger than G.C.H., but known to be relatively consistent with the ordinary axioms of set theory.

THEOREM 7.9 (Jensen). *Assuming $V = L$, 7.8 also applies when κ is not regular.*

The status of 7.9 assuming only G.C.H. remains open. Much more is known about two-cardinal and related problems. See [3] for some of these results and references to work of MacDowell and Specker, Fuhrken, Keisler, Morley, Silver, the author, and others. A recent result is [13]. Very recently Jensen, assuming $V = L$, has obtained results concerning pairs (k^{++}, k) , (k^{+++}, k) , etc.

There are several books on model theory. Bell and Slomson [2] is a good shorter treatment. Chang and Keisler [3] is by far the most comprehensive treatment to appear. (Incidentally, most references we have omitted can be found in [3].) A. Robinson ([12] and others) is excellent for applications of model theory to algebra. Of the many important topics which we have not dealt with, the one which is closest to those which we have discussed is the application of the partition theorems of Ramsey and Erdős-Rado to model theory — by Ehrenfeucht-Mostowski, Morley, and many others. Naturally, a discussion of this topic can be found in [3].

Part of the author's work was supported by National Science Foundation Grant No. NSF-GP-8746.

References

1. J. Ax and S. Kochen, Diophantine problems over local fields I, II, III, *Amer. J. Math.*, 87, (1965) pp. 605–630, 631–468; *Ann. Math.*, vol. 83, pp. 437–456.
2. J. Bell and A. Slomson, *Models and Ultraproducts: an Introduction*, North Holland, Amsterdam, 1969.
3. C. C. Chang and H. J. Keisler, *Model Theory* (to appear).
4. B. Jonsson, Homogeneous universal relational systems, *Math. Scand.*, 8 (1960) 137–142.
5. C. Karp, *Languages with expressions of infinite length*, North Holland, Amsterdam, 1964.
6. H. J. Keisler, Some model-theoretic results for ω -logic, *Israel. J. Math.*, 4(1965) 249–261.
7. ———, Logic with the quantifier “there exist uncountably many”, *Ann. of Math. Logic*, 1 (1970) 1–94.
8. E. G. K. Lopez-Escobar, On defining well-orderings, *Fund. Math.*, 57 (1965) 253–272.
9. M. Morley, Categoricity in power, *Trans. Amer. Math. Soc.*, 114 (1965) 514–538.
10. ———, Countable models of \aleph_1 -categorical theories, *Israel J. Math.*, 5 (1967) 65–72.
11. M. Morley and R. Vaught, Homogeneous universal models, *Math. Scand.*, 11 (1962) 37–57.
12. A. Robinson, *Complete Theories*, North Holland, Amsterdam, 1956.
13. J. Schmerl and S. Shelah, On models with orderings, *Notices, Amer. Math. Soc.*, 17 (1970) 294.
14. A. Tarski, Some notions and methods on the borderline of algebra and metamathematics, *Proc. Int. Cong. Math.*, Providence, 1950, 1952, pp. 705–720.
15. ———, *A Decision Method for Algebra and Geometry*, 2nd ed., University of California Press, Berkeley and Los Angeles, 1951.
16. B. L. van der Waerden, *Modern Algebra*, vol. 1, Ungar, New York, 1964.
17. R. Vaught, Denumerable models of complete theories. Infinitistic methods, *Proc. Symposium in Foundations Math.*, Warsaw 1959, New York, 1961, pp. 303–321.
18. J. Baldwin and A. Lachlan, On strongly minimal sets, *J. Symb. Logic*, 36 (1971) 79–96.

WHAT IS NONSTANDARD ANALYSIS?

W. A. J. LUXEMBURG, California Institute of Technology

1. Introduction. The subject referred to in the title with which we shall deal may seem perhaps at first sight to be far removed from the general topic "The Foundations of Mathematics" of the Symposium. This relatively new field which was created by Abraham Robinson (see [7]) may be looked upon, however, as a major contribution to the foundations of analysis. Furthermore, it is another splendid example of an application of mathematical logic.

The development of mathematical analysis by using infinitely small and infinitely large numbers has been a subject of constant interest and controversy in the history of mathematics. Going back in history we discover that Leibniz was one of the strongest advocates of a method involving infinitely small and infinitely large numbers in the early stages of the development of the calculus. The reason why the theory of infinitesimals gradually fell into disrepute and was replaced later by the ε, δ -method must be sought in the fact that neither Leibniz nor his successors were able to state with sufficient precision just what rules were supposed to govern their system of infinitely large and infinitely small numbers. Although Leibniz stated the principle that what holds for the finite numbers should also hold for the numbers in the extended system, which includes the infinitely small and infinitely large numbers, it is not at all clear in his writings what sort of laws about numbers his principle was supposed to apply to.

It was Abraham Robinson's recent discovery, mentioned above, that the notions of model theory can clarify the notions of infinitely small and infinitely large. Robinson shows that mathematical analysis can be developed by imbedding the real number system R in a proper extension *R of R which possesses in a certain sense the same properties as R . It is well known that such an extension *R must be non-Archimedean and this is the fact that enabled Robinson to define in *R the infinitely small and infinitely large numbers whose existence was taken for granted by Leibniz and his followers. From the well-known result that there exist systems of axioms for the real number system which are categorical, that is, determine the real number systems uniquely up to isomorphism it may seem at first very paradoxical that such systems *R exist. This sort of paradox has been one of the main sources of the condemnation of the theory of infinitesimals and infinitely large numbers as a tool in analysis. The paradox vanishes completely, however, if we follow Robinson's idea to restrict the statement "the same properties" to a specified collection of properties of R which can be formulated in a specified formal language with the appropriate interpretation in R as well as in *R , and in which the classical isomorphism theorem for the real number system cannot be formulated. Of course it is at this

point that model theory comes into play which by means of the compactness principle guarantees the existence of such systems *R .

There is, however, another way to establish the existence of *R . This method is known as the construction of models in the form of ultraproducts. It has the advantage that it can be developed within the framework of axiomatic set theory. We shall follow this procedure here. Sections 2, 3 and 4 are entirely devoted to a discussion of the existence of *R . In our approach we follow very closely the development as given by Abraham Robinson and Elias Zakon in their paper entitled *A set-theoretical characterization of enlargements* and which appeared in [6]. In the remaining six sections it is illustrated by means of examples in which sense the theory of infinitely small and infinitely large numbers can be used as a tool in analysis. The topics which were selected for this purpose include the theory of limits, Euler's product formula for the sine, and the existence of functions which are not measurable in the sense of Lebesgue.

The ideas of nonstandard analysis were subsequently successfully applied to other branches of mathematics. These developments are not taken up here as they are beyond the scope of the present introductory paper. But we like to refer the interested reader, who for instance would like to know with what great success this method was used by A. Robinson and A. Bernstein to solve the invariant subspace problem for a certain class of bounded operators on a Hilbert space, to Robinson's book [8] and the papers [1], [2] and [15]. Furthermore, we would like to draw the readers' attention to reference [6] which is the *Proceedings of the International Symposium on Nonstandard Analysis*, which was held at the California Institute of Technology in 1967. Its contents, consisting of more than twenty papers, gives the latest developments in this field.

Finally, the author would like to state that the present paper is mainly expository in nature. It is particularly directed to those mathematicians who would like to get acquainted with this new tool in analysis. We do hope, however, that also the specialists in the field will find something new and of interest in this paper.

2. Definition of the structure \hat{R} and some of its properties. The earlier version of nonstandard analysis (see [7] and [3]) rests on the formulation of the properties of R which can be formulated in a first order language, which means briefly that quantification in the formal language is permitted only on variables ranging over real numbers. One need not go far in analysis, however, to realize the need for a richer language in which statements containing expressions such as for example "For all nonempty sets of natural numbers..." or "There exists a continuous function..." can be formulated. In this connection it is also good to observe that even some of the axioms of the real number system are outside the language of the lower predicate calculus. For example, Dedekind's completion axiom involving quantification with respect to ordered pairs of sets (Dedekind cuts) is such an axiom. In order to cope with this difficulty we shall use the framework of axiomatic set

theory in terms of which the theory of real numbers can be developed. The formal language will be a lower order language whose constants will range over sets and numbers. We shall now present this development here in some detail. We shall assume that the reader is familiar with the elements of naive set theory and with some of the definitions and results concerning the lower predicate calculus.

Let R denote as usual the set of real numbers. Then we define inductively the sets $R_0 = R$ and $R_{n+1} = P(\bigcup_{k=0}^n R_k)$ ($n = 0, 1, 2, \dots$), where $P(X)$ denotes the set of all subsets of X . The union of all the sets R_n , $\bigcup_{n \geq 0} R_n$ is called the **superstructure** on R and will be denoted by \hat{R} . The elements of \hat{R} are called the **entities** of the superstructure \hat{R} . The elements of $R_0 = R$, that is the real numbers, on which the superstructure is based are sometimes also referred to as the **individuals** of \hat{R} .

We shall assume that an ordered pair (a, b) is defined in the sense of Kuratowski by $(a, b) = \{\{a\}, \{a, b\}\}$ and that n -tuples (a_1, \dots, a_n) are defined inductively by $(a) = a$, $(a_1, \dots, a_n) = ((a_1, \dots, a_{n-1}), a_n)$. Then it follows immediately that relations defined as sets of n -tuples ($n = 1, 2, \dots$) are all entities of \hat{R} . Since the algebraic operations of R can be defined in terms of three place relations as follows: $ab = c$ if and only if $(a, b, c) \in P \in \hat{R}$ and $a + b = c$ if and only if $(a, b, c) \in S \in \hat{R}$ and the order relation is a binary relation it follows that the axioms and the properties of R can be expressed in terms of certain entities of \hat{R} . The remaining part of this section will now be devoted to making this more precise.

The entities of $R_n - R_{n-1}$ ($n \geq 1$) are called of rank n in \hat{R} . The individuals are given the rank 0. The reader should observe that by means of this definition, the empty set gets assigned rank 1. If $a \in \hat{R}$ is not empty, then the rank of a is the smallest natural number n such that $a \in R_n$. It is also easy to see that if $a_1, \dots, a_n \in \hat{R}$, then $\text{rank}(a_1, \dots, a_n) = \max(\text{rank } a_1, \dots, \text{rank } a_n) + 2n$.

Some minor set-theoretical properties of \hat{R} are collected, for later references, in the following lemma.

LEMMA 2.1. (i) $R_p \subset R_n$ for all $n \geq p \geq 1$.
 (ii) $\bigcup_{k=0}^n R_k = R_0 \cup R_n$ for all $n \geq 1$.
 (iii) $R_k \in R_{n+1}$ for all $0 \leq k \leq n$ and for all $n \geq 0$.
 (iv) If $x \in y \in R_n$ ($n \geq 1$), then $x \in R_0 \cup R_{n-1}$.
 (v) If $(x_1, \dots, x_n) \in y \in R_p$ ($p \geq 1$), then $x_1, \dots, x_n \in R_0 \cup R_{p-1}$. In particular, if an entity $\Phi \in \hat{R}$ is a binary relation, then its domain, $\text{dom } \Phi = \{x : (\exists y)(x, y) \in \Phi\} \in \hat{R}$, and its range, $\text{ran } \Phi = \{y : (\exists x)(x, y) \in \Phi\} \in \hat{R}$.

Proof. (i) If $x \in R_p$, then $x \subset \bigcup_{k=0}^{p-1} R_k$, and so $x \subset \bigcup_{k=0}^q R_k$ for all $q \geq p-1$. Hence, $x \in P(\bigcup_{k=0}^q R_k) = R_{q+1}$ for all $q+1 \geq p$.

(ii) For $n \geq 1$, $R_n \subset R_{n+1}$, and so since R_0 is disjoint from all R_n ($n \geq 1$) it follows that for all $n \geq 1$ we have $\bigcup_{k=0}^n R_k = R_0 \cup R_n$.

(iii) Since by (ii) we have that $R_k \subset R_0 \cup R_n$ ($0 \leq k \leq n$) we obtain that $R_k \in P(R_0 \cup R_n) = R_{n+1}$.

(iv) If $y \in R_n$, ($n \geq 1$), then $y \subset R_0 \cup R_{n-1}$, and so $x \in y$ implies that $x \in R_0 \cup R_{n-1}$.

(v) If $(x_1, \dots, x_n) \in y \in R_p$ ($p \geq 1$), then $(x_1, \dots, x_n) \in R_0 \cup R_{p-1}$. Hence, $\{\{x_1\}, \{x_1, (x_2, \dots, x_n)\}\} \in R_0 \cup R_{p-1} = R_0 \cup P(R_0 \cup R_{p-2})$ implies $x_1 \in R_0 \cup R_{p-2} \subset R_0 \cup R_{p-1}$, and similarly for the entities x_2, \dots, x_n .

The formal language will now be introduced.

The **atomic symbols** of L are: (i) The **connectives** $\wedge, \vee, \Rightarrow, \Leftrightarrow, \neg$, for “and”, “or”, “implies”, “if and only if”, “not” respectively. (ii) The **variables**, a countably infinite sequence usually denoted by x, y, \dots with or without subscripts. (iii) The **quantifiers** $(\exists \cdot)$ -existential, and $(\forall \cdot)$ -universal. (iv) Brackets $[\]$, used for grouping formulas as usual in mathematics. (v) The **basic predicate**, \in read “member of” with one open place to the left and to the right of it. (vi) **Extra logical constants** (briefly, constants). This is a set of symbols of which there are enough to be put in one-to-one correspondence with the entities of whatever structure may be under consideration. This set of constants is usually infinite but fixed. Furthermore, constants are usually denoted by Roman letters with or without subscripts from the beginning of the alphabet, and other symbols such as the numerals $0, 1, 2, \dots$.

We shall now assume that the set of constants of L is brought in one-to-one correspondence with all the entities of the structure \hat{R} and we shall from now on identify the constants of L with the entities of \hat{R} so that \hat{R} is part of L . If such an identification has been established, then we refer to \hat{R} as an L -structure.

The interpretation of the basic predicate \in of L in \hat{R} will be the membership relation of axiomatic set theory.

From the atomic formulas $\alpha \in \beta$, where the symbols α and β may denote constants and variables, the well-formed formulas (wff) are obtained in successive stages by applying the connectives and quantifiers. At the same time brackets are introduced in such a way that the formation of the formula can be unambiguously determined. More precisely, if V is an atomic formula, then $[V]$ is a wff, if V, W are wff, then $[V \wedge W], [V \vee W], [\neg V], [V \Rightarrow W], [V \Leftrightarrow W]$ are wff; and if V is a wff, then $[(\forall x)V]$ and $[(\exists x)(V)]$ are wff, where x denotes an arbitrary variable, provided x does not already appear in V under the sign of a quantifier. Furthermore, we shall adhere to the terminology that in $[(\forall x)V]$ and $[(\exists x)V]$, V is called the **scope** of the quantifier and in all the wff which can be obtained from these by the further repeated applications of connectives and quantifiers. A variable x is called *free* in a wff V if x is not in $(\exists x)$ or $(\forall x)$ or in the scope of a quantifier in V . A wff is called a **sentence** if every variable is in the scope of a quantifier, otherwise it is called a **predicate**. A wff V in L is said to be in **prenex normal form**, if in the formation of V from atomic formulas the quantifiers are applied after the connectives, that is, if the connectives are in the scope of *all* quantifiers. In symbols, $V = (qx_n) \dots (qx_1)W$, where $(q \cdot)$ denotes either $(\exists \cdot)$ or $(\forall \cdot)$ and where W is a wff without quantifiers, is a wff in prenex normal form. One of the basic results of the lower predicate calculus states that every wff is equivalent to a wff which is in prenex normal form (see [8], p. 10).

For our purpose we shall only consider those wff of L which have the property that all quantifiers are of the form " $(\forall x)[[x \in A] \Rightarrow \dots]$ " and " $(\exists x)[[x \in A] \wedge \dots]$ " where A is an entity of \hat{R} and which are called the **admissible** wff. Thus a wff is admissible whenever the domain of every quantifier occurring in it is a specific entity of \hat{R} . The set of admissible wff of L will be denoted by $K = K(L)$ and the subset of K of all admissible sentences which hold in \hat{R} will be denoted by $K_0 = K_0(L)$.

At this point the reader should do well to observe that all statements in analysis dealing with numbers, sets of numbers, relations between numbers, relations between sets and numbers, and so on, and which hold in R can be expressed as admissible sentences of L which are in K_0 . For instance, the sentence of K_0

$$(\forall a)(\forall b)(\forall c)[a, b, c \in R] \Rightarrow [P(a, b, c) \Rightarrow P(b, a, c)]$$

expresses that R is commutative (P is the constant denoting the three place relation of multiplication).

Any $*L$ -structure $*(\hat{R})$ in which the L -structure \hat{R} can be properly imbedded and for which all admissible sentences of \hat{R} which hold in \hat{R} with appropriate interpretation of the symbols in $*(\hat{R})$ also hold in $*(\hat{R})$ will be called a **higher order nonstandard model** of \hat{R} . In that case, it turns out that the set $*R$ of individuals of $*(\hat{R})$ is a totally ordered field of which R is a proper subfield. But $*(\hat{R})$ is not the superstructure determined by $*R$. In fact, if $A = P(R)$ is the constant which denotes the entity of \hat{R} of all subsets of R , then under the imbedding of \hat{R} in $*(\hat{R})$ this constant will not denote the set of *all* subsets of $*R$ as might be expected at first, but only a subsystem of the power set of $*R$, and so on. How this all will come about will be explained in detail in the next section.

3. Models of \hat{R} that are ultrapowers. We begin by recalling some definitions and elementary results from the theory of filters.

Let I denote a nonempty set. By a **filter** over I we mean a nonempty set \mathcal{F} of subsets of I such that the empty set $\emptyset \notin \mathcal{F}$, \mathcal{F} is closed under **finite** intersections, and $F \subset G$ and $F \in \mathcal{F}$ implies $G \in \mathcal{F}$. In particular $\mathcal{F} \neq \emptyset$ implies that $I \in \mathcal{F}$. A filter \mathcal{F}_1 is called **finer** than a filter \mathcal{F}_2 ($\mathcal{F}_2 \leq \mathcal{F}_1$) whenever $F \in \mathcal{F}_2$ implies $F \in \mathcal{F}_1$. This relation orders the set of all filters over I and the filter $\{I\}$ is its smallest element. A filter \mathcal{F} is called an **ultrafilter** whenever it is not properly contained in any other filter, that is, the ultrafilters are the maximal elements of the ordered set of filters. Concerning ultrafilters we have the following important characterization. A filter \mathcal{F} is an ultrafilter if and only if for every $F \subset I$ either $F \in \mathcal{F}$ or $I - F \in \mathcal{F}$. The latter statement is easily seen to be equivalent to: If

$$\bigcup_{i=1}^n F_i \in \mathcal{F} (F_i \subset I, i = 1, 2, \dots, n),$$

then $F_i \in \mathcal{F}$ for at least one index i , and so, is itself a characterization of the concept of an ultrafilter.

A filter \mathcal{F} is called δ -incomplete, whenever there exists a sequence $F_n \in \mathcal{F}$ ($n = 1, 2, \dots$) such that $\bigcap_{n=1}^{\infty} F_n \notin \mathcal{F}$, and a filter \mathcal{F} is called δ -complete whenever it is not δ -incomplete. A filter \mathcal{F} is called **free** whenever $\bigcap (F: F \in \mathcal{F}) = \emptyset$. It is not known whether δ -complete free ultrafilters exist. This problem is known as Ulam's measure problem. It is easy to see, however, that a δ -incomplete ultrafilter is free. It follows from the following simple result, the proof of which we leave to the reader as an exercise.

An ultrafilter \mathcal{U} is δ -incomplete if and only if there exists a countable partition $\{I_n: n = 1, 2, \dots\}$ of the set I over which \mathcal{U} is defined such that $I_n \notin \mathcal{U}$ for all $n = 1, 2, \dots$.

From this result in conjunction with Zorn's lemma it follows now also easily that on every infinite set there exist plenty of δ -incomplete ultrafilters. For further information on filters we refer the reader to the paper of the author: *A general theory of monads*; which appeared in [6].

We shall now turn to a description of a structure which is an ultrapower of \hat{R} .

Let I be an infinite set, let \mathcal{U} be a δ -incomplete ultrafilter of subsets of I and let $\{I_n: n = 1, 2, \dots\}$ be a countable partition of I satisfying $I_n \notin \mathcal{U}$ for all $n = 1, 2, \dots$ which will be kept fixed.

By \hat{R}^I we denote as usual the set of all mappings of I into \hat{R} . There exists a natural imbedding $a \rightarrow *a$ of \hat{R} into \hat{R}^I defined by $*a(i) = a$ for all $i \in I$, that is \hat{R} is identified in \hat{R}^I by the constant mappings. The undefined basic predicates “=” and “ ϵ ” of \hat{R} can be extended to \hat{R}^I by means of the following \mathcal{U} -dependent definitions.

DEFINITION 3.1. *If $a, b \in \hat{R}^I$, then $a =_{\mathcal{U}} b$ if and only if $\{i: a(i) = b(i)\} \in \mathcal{U}$, and $a \epsilon_{\mathcal{U}} b$ if and only if $\{i: a(i) \epsilon b(i)\} \in \mathcal{U}$.*

Since it is an immediate consequence of $I \in \mathcal{U}$ that if $a, b \in \hat{R}$, then $a = b$ if and only if $*a =_{\mathcal{U}} *b$, and $a \epsilon b$ if and only if $*a \epsilon_{\mathcal{U}} *b$ it follows that the relations “ $=_{\mathcal{U}}$ ” and “ $\epsilon_{\mathcal{U}}$ ” are \mathcal{U} -extensions of “=” and “ ϵ ” of \hat{R} . For the sake of simplicity we shall from now on retain the original notation “=” for “ $=_{\mathcal{U}}$ ” and “ ϵ ” for “ $\epsilon_{\mathcal{U}}$ ”.

In order to justify the definition we are going to show that for all $a, b \in \hat{R}^I$ either $a = b$ or not ($a = b$) ($a \neq b$) holds, and $a \epsilon b$ or not ($a \epsilon b$) ($a \not\epsilon b$) holds. Since the proof for both cases is the same we shall only verify it for “=” . If $a, b \in \hat{R}^I$, then we set

$$U_1 = \{i: a(i) = b(i)\} \text{ and } U_2 = \{i: a(i) \neq b(i)\}.$$

From $U_1 \cup U_2 = I \in \mathcal{U}$ it follows from the basic property of an ultrafilter that either $U_1 \in \mathcal{U}$ and $U_2 \notin \mathcal{U}$ or $U_1 \notin \mathcal{U}$ and $U_2 \in \mathcal{U}$, that is, by Definition 3.1., either $a = b$ or not ($a = b$) holds.

Having justified the definition we can justify further the suggestion that the relations “=”, “ \in ” in \hat{R}^I behave like equality and membership of set theory. Since the individuals of \hat{R} are without members but different from \emptyset , that is, set theory in \hat{R} is based on a set of so-called urelements, equality of sets in terms of \in should read “ $a = b$ ” if and only if $a \in c$ and $b \in c$ for all $c \in \hat{R}^I$. But this can now be immediately verified by observing that if $a = b$ and $a \in c$, then

$$U_1 = \{i: a(i) = b(i)\} \in \mathcal{U} \text{ and } U_2 = \{i: a(i) \in c(i)\} \in \mathcal{U}$$

implies by the filter properties that $U_1 \cap U_2 \in \mathcal{U}$, and so $i \in U_1 \cap U_2$ implies that $b(i) \in c(i)$, that is, $b \in c$. Conversely, we have that if $a, b \in \hat{R}^I$, then $a \in \{x: x = a \text{ and } x \in \hat{R}^I\} = \{a\}$, implies that $b \in \{a\}$, that is $b = a$. That the relation of equality, as defined in Definition 3.1, is an equivalence relation is immediately clear. That it satisfies the rule of substitution in \in , namely,

$$(\forall a)(\forall b)(\forall c)(\forall d) [[a \in b] \wedge [a = c] \wedge [b = d]] \Rightarrow [c \in d]$$

can be verified in the same way by using the properties of \mathcal{U} .

Continuing this process we can show, by using the basic properties of \mathcal{U} , that one by one the statements which hold in \hat{R} hold in \hat{R}^I under the defined interpretation of the basic predicates. We shall of course not follow this procedure but present in a general fashion that a certain substructure of \hat{R}^I has the same properties as \hat{R} .

For this purpose we shall assume that the elements of \hat{R}^I are identified in a one-to-one manner with the constants of a formal language $*L$. Furthermore, $*L$ is assumed to have two basic predicates “=” (equality) and “ \in ” (membership) which are identified with the corresponding relations of \hat{R}^I . Thus we obtain an $*L$ -structure \hat{R}^I whose set of true sentences depends on \mathcal{U} . A certain substructure of our $*L$ -structure will be singled out which we shall show to satisfy, in a certain sense, the sentences of K_0 .

In the following lemma, however, we shall first list for later reference, some of the basic properties of the imbedding $a \rightarrow *a$ of \hat{R} into \hat{R}^I .

LEMMA 3.2. (i) $*\emptyset = \emptyset$.

(ii) If $a, b \in \hat{R}$, then $a \subset b$ implies $*a \subset *b$.

(iii) If $a, b \in \hat{R}$, then $a \in b$ if and only if $*a \in *b$.

(iv) For all $a \in \hat{R}$ we have $*\{a\} = \{*a\}$.

(v) If $a_1, \dots, a_n \in \hat{R}$, then $*(\bigcup_{i=1}^n a_i) = \bigcup_{i=1}^n *a_i$, $*(\bigcap_{i=1}^n a_i) = \bigcap_{i=1}^n *a_i$, $*\{a_1, \dots, a_n\} = \{*a_1, \dots, *a_n\}$, $*(a_1, \dots, a_n) = (*a_1, \dots, *a_n)$, and $*(a_1 \times \dots \times a_n) = *a_1 \times \dots \times *a_n$.

(vi) For all $a, b \in \hat{R}$ we have $*(a - b) = *a - *b$.

(vii) If $b \in \hat{R}$ is a binary relation, then $*(\text{dom } b) = \text{dom } *b$, $*(\text{ran } b) = \text{ran } *b$, and for all $a \in \hat{R}$ we have

$$*(b(a)) = *\{y: (\exists x)(x \in a \wedge (x, y) \in b)\} = *b(*a) = \{y: (\exists x)(x \in *a \wedge (x, y) \in *b)\}.$$

Proof. We shall only prove (vi) since the proofs of the other statements are similar. For these proofs we refer the reader to the proofs of Theorems 7.1 and 7.7. of [3].

(vi) If $c \in *(a-b)$, then $U_1 = \{i: c(i) \in a-b\} \in \mathcal{U}$ implies, using $U_1 \subset U_2 = \{i: c(i) \in a\} \in \mathcal{U}$ that $c \in *a$ and using $U_1 \subset U_3 = \{i: c(i) \notin b\} \in \mathcal{U}$ which, since \mathcal{U} is an ultrafilter, is equivalent to $\{i: c(i) \in b\} \notin \mathcal{U}$ that $c \notin *b$, and so $c \in *a - *b$. For the converse reverse the steps.

DEFINITION 3.3. *An entity a of the $*L$ -structure \hat{R}^I is called internal whenever there exists a natural number $n \geq 0$ such that $a \in *R_n$. An internal entity a is called a standard entity whenever there exists an entity $b \in \hat{R}$ such that $a = *b$. All entities which are not internal are called external.*

*The set $\bigcup_{n \geq 0} *R_n$ of all internal entities is called the ultrapower of \hat{R} with respect to the ultrafilter \mathcal{U} and will be denoted by $*(\hat{R})$.*

The \mathcal{U} -ultrapower of \hat{R} is usually denoted by \mathcal{U} -prod \hat{R} but we shall not employ this notation in this paper.

Observe that the mapping $a \rightarrow *a$ of \hat{R} into \hat{R}^I imbeds \hat{R} into the substructure $*(\hat{R})$ of \hat{R}^I .

The notion of rank extends immediately to the internal entities. An internal entity $a \in *(\hat{R})$ is said to be of rank n ($n \geq 1$) whenever $a \in *R_n - *R_{n+1}$; and the entities of $*R = *R_0$ are said to be of rank 0. The entities of rank 0 are also referred to as the **individuals** of $*(\hat{R})$. Again, by means of this definition the empty set $*\emptyset$ has rank 1. The rank of an internal entity can be further specified. If a is non-empty and internal, then $a \in *R_p$ for some $p \geq 0$, and so, by Definition 3.1, we have that $U = \{i: a(i) \in R_p\} \in \mathcal{U}$. Then $\bigcup_{k=0}^p \{i: \text{rank } s(i) = k\} = U \in \mathcal{U}$ implies, using the fact that \mathcal{U} is an ultrafilter, that there exists exactly one index n such that $0 \leq n \leq p$ and $U_1 = \{i: \text{rank } s(i) = n\} \in \mathcal{U}$. Then for all $i \in U_1 \in \mathcal{U}$ we have $a(i) \in R_n - R_{n-1}$, and so $a \in *(R_n - R_{n-1}) = *R_n - *R_{n-1}$ (Lemma 3.2(vi)), that is, rank $a = n$.

If $a = *b$, $b \in \hat{R}$, is a standard entity of $*(\hat{R})$, then its rank remains unchanged.

At this point it seems natural to ask the question whether there are internal entities which are not standard. Fortunately, the answer to this question is affirmative and as we shall see in the following theorem it is a consequence of the hypothesis that the ultrafilter \mathcal{U} is δ -incomplete, a hypothesis which we have not used so far.

THEOREM 3.5. *There exist internal entities which are not standard. In fact, if $a \in \hat{R}$ is an entity which has infinitely many elements, then there exists an entity $b \in *a$ such that b is not standard.*

Proof. Since a is an infinite set there exists a sequence $\{b_n: n = 1, 2, \dots\}$ of elements of a such that $b_n \neq b_m$ for all $n, m = 1, 2, \dots$ and $n \neq m$. Let b be the mapping of I into a such that $b(i) = b_n$ for all $i \in I_n$ ($n = 1, 2, \dots$). Then $b \in *a$ but b is not equal to any standard element of $*(\hat{R})$, and the proof is complete.

The internal entities, defined to be the elements of the special standard sets $*R_n$, can also be characterized as follows. *An entity a is internal if and only if a is an element of a standard entity.* In order to see this we need only to show that if $a \in *b$, $b \in \hat{R}$, then a is internal. Now from $b \in \hat{R}$ it follows that $b \in R_n$ for some n which implies that $b \subset R_0 \cup R_{n-1}$, and so, by Lemma 3.2(v), $a \in *b \subset *R_0 \cup *R_{n-1}$ implies $a \in *R_0 \cup *R_{n-1}$ which shows that a is internal. In view of Theorem 3.5, we may ask the question, what about the nature of the entities which are elements of internal entities? The answer is that they are internal, as the following theorem shows. The converse, however, is not true. In fact, we shall see later in Section 5 that a set of internal entities need not be internal.

THEOREM 3.6. *If $a \in b \in *R_n$ ($n \geq 1$), then $a \in *R_{n-1}$, that is, the elements of an internal entity are internal.*

Proof. From $b \in *R_n$ it follows that $U = \{i: b(i) \subset R_0 \cup R_{n-1}\} = \{i: b(i) \in R_n\} \in \mathcal{U}$, and so for all $i \in U$ we have $a(i) \in R_0 \cup R_{n-1}$. Hence, by Lemma 3.2(v) and Definition 3.1, $a \in *(R_0 \cup R_{n-1}) = *R_0 \cup *R_{n-1}$, and the proof is finished.

As in the case of the L -structure \hat{R} we shall call an $*L$ -wff **admissible** whenever all the quantifiers occurring in it are of the form “ $(\forall x) [[x \in a] \Rightarrow \dots]$ ” and “ $(\exists x) [[x \in a] \wedge \dots]$ ”, where a is a constant denoting an entity of \hat{R}^I .

*An admissible wff of $*L$ is called internal whenever all the constants occurring in it denote internal entities. An admissible wff of $*L$ is called standard whenever all the constants occurring in it denote standard entities.* Thus a standard wff is internal.

The set of all **internal sentences** of $*L$ will be denoted by $*K = *K(*L)$, and the subset of all **internal sentences** which hold in $*(\hat{R})$ will be denoted by $*K_0 = *K_0(*L)$.

If V is an admissible wff of L , then its **$*$ -transform** $*V$ is defined to be that **standard wff** of $*L$ which is obtained from V by replacing in V all the constants, say, a_1, \dots, a_p , occurring in it, by $*a_1, \dots, *a_p$ but leaving the variables and bracketing unchanged.

We shall now prove that the $*$ -imbedding has the following important property.

THEOREM 3.7. *Let $V = V(x_1, \dots, x_p)$ be an admissible L -wff with the free variables x_1, \dots, x_p , and let $A = \{(x_1, \dots, x_p): (x_1, \dots, x_p) \in a \text{ and } V(x_1, \dots, x_p)\}$, where a is an arbitrary entity of \hat{R} . Then $A \in \hat{R}$ and*

$$*A = \{(y_1, \dots, y_p): (y_1, \dots, y_p) \in *a \text{ and } *V(y_1, \dots, y_p)\}.$$

Proof. That $A \in \hat{R}$ is trivial. If $V = V(x_1, \dots, x_p, a_1, \dots, a_q)$ is atomic, that is, V has the form $(x_1, \dots, x_p, a_1, \dots, a_{q-1}) \in a_q$ or $(x_1, \dots, x_{p-1}, a_1, \dots, a_q) \in x_p$ with possible permutation of the variables, then the result follows immediately from

Definition 3.1. In order to show that the result holds for all wff V of L without quantifiers we have to show that if it holds for two such wff V and W , then it also holds for $[V \wedge W]$ and $[\neg V]$. As is well known this will take care of all the logical connectives. Assume that $*A = \{(x_1, \dots, x_p): (x_1, \dots, x_p) \in *a \text{ and } *V(x_1, \dots, x_p)\}$, then we have to show that

$$*B = \{(x_1, \dots, x_p): (x_1, \dots, x_p) \in *a \text{ and } \neg *V(x_1, \dots, x_p)\},$$

where $B = a - A$. Since, by Lemma 3.2(vi), $*B = *a - *A$ the result follows. Assume now that $V = V(x_1, \dots, x_p, y_1, \dots, y_q)$ and $W = W(x_1, \dots, x_p, z_1, \dots, z_r)$ be two L -wff without quantifiers for which the result holds, and let

$$A = \{(x_1, \dots, x_p, y_1, \dots, y_q, z_1, \dots, z_r): (x_1, \dots, x_p, y_1, \dots, y_q, z_1, \dots, z_r) \in a \text{ and } [V \wedge W]\}.$$

Then $A = \{(x_1, \dots, z_r): (x_1, \dots, z_r) \in a \text{ and } V\} \cap \{(x_1, \dots, z_r): (x_1, \dots, z_r) \in a \text{ and } W\}$ implies, by Lemma 3.2(v), that $*A = *\{\dots\} \cap *\{\dots\} = \{(x_1, \dots, z_r): (x_1, \dots, z_r) \in *a \text{ and } [V \wedge W]\}$, and so the result holds for all wff without quantifiers.

For admissible wff with quantifiers we shall use induction on the number n of quantifiers. For $n = 0$ the result was shown above. Assume now that the result holds for all admissible wff with less than or equal n quantifiers. Let V be an admissible wff with $(n+1)$ -quantifiers which is written in its prenex normal form $(qx_{n+1}) \dots (qx_1)W(x_1, \dots, x_{n+1}, y_1, \dots, y_q)$, where W has no quantifiers and y_1, \dots, y_q are the free variables occurring in V . Without loss of generality we may assume that (qx_{n+1}) is the existential quantifier $(\exists x_{n+1})$ otherwise we consider not V . Let b denote the domain of $(\exists x_{n+1})$. Then since V is admissible, $b \in \hat{R}$. Let

$$B = \{((y_1, \dots, y_p), x_{n+1}): ((y_1, \dots, y_p), x_{n+1}) \in a \times b \text{ and } (qx_n) \dots (qx_1)W\},$$

where $a \in \hat{R}$. Then, by the induction hypothesis and Lemma 3.2(v), we obtain that

$$*B = \{((y_1, \dots, y_p), x_{n+1}): ((y_1, \dots, y_p), x_{n+1}) \in *a \times *b \text{ and } (qx_n) \dots (qx_1)*W\}.$$

The domain of the binary relation B is the set

$$\begin{aligned} A &= \{(y_1, \dots, y_p): (y_1, \dots, y_p) \in a \text{ and} \\ &\quad (\exists x_{n+1})(x_{n+1} \in b \wedge (qx_n) \dots (qx_1)W)\} \\ &= \{(y_1, \dots, y_p): (y_1, \dots, y_p) \in a \text{ and } V(y_1, \dots, y_p)\}. \end{aligned}$$

The domain of the binary relation $*B$ is, however, the set

$$\begin{aligned} &\{(y_1, \dots, y_p): (y_1, \dots, y_p) \in *a \text{ and } (\exists x_{n+1})(x_{n+1} \in *b \wedge (qx_n) \dots (qx_1)*W)\} \\ &= \{(y_1, \dots, y_p): (y_1, \dots, y_p) \in *a \text{ and } *V\}. \end{aligned}$$

Then, by Lemma 3.2(vii), we obtain the desired result that

$$*A = \{(y_1, \dots, y_p) : (y_1, \dots, y_p) \in *a \text{ and } *V\},$$

and the proof is finished.

We are now in a position to prove the *Fundamental Theorem* about ultrapowers which we shall refer to throughout the rest of the paper by F.T.

THEOREM 3.8. *$*(\hat{R})$ is a higher order nonstandard model of \hat{R} , that is, an admissible sentence V of $K(L)$ holds in \hat{R} if and only if $*V$ holds in $*(\hat{R})$, and \hat{R} is properly imbedded in $*(\hat{R})$.*

Proof. Theorem 3.5 tells us that the imbedding $a \rightarrow *a$ of \hat{R} into $*(\hat{R})$ is proper. We have to show that if $V \in K(L)$, then $V \in K_0$ if and only if $*V \in *K_0$. If V has no quantifiers, then it follows immediately from Definition 3.1. Assume that $V \in K$ has the prenex normal form $V = (qx_n) \dots (qx_1)W$, where W has no quantifiers. There is no loss in generality to assume that (qx_n) is the existential quantifier $(\exists x_n)$. Then $V \in K_0(L)$ is equivalent to “the set $A = \{x_n : x_n \in a \text{ and } (qx_{n-1}) \dots (qx_1)W\} \neq \emptyset$, where a is the domain of $(\exists x_n)$. Then, by Theorem 3.7 and Lemma 3.2(i), we see that $A \neq \emptyset$ is equivalent to $*A = \{x_n : x_n \in *a \text{ and } (qx_{n-1}) \dots (qx_1)*W\} \neq *\emptyset$ which itself is equivalent to $*V \in *K_0$, and the proof is finished.

An important aspect of the method of nonstandard analysis is to use the F.T. repeatedly to transform the true statements of \hat{R} into true statements about the internal entities of $*(\hat{R})$. To illustrate this we shall give a number of examples dealing with the set theory of \hat{R} .

EXAMPLES 3.9. (i). The individuals of \hat{R} are the “urelements” of the set theory of \hat{R} in the sense that although they are different from the empty set \emptyset there are no entities of \hat{R} which are elements of individuals. This true statement can be expressed by the following infinite list of sentences of K_0 .

$$(\forall x)(\forall y)[x \in R] \wedge [y \in R_n] \Rightarrow [\neg y \in x], \quad n = 0, 1, 2, \dots$$

From the F.T. we conclude that $*K_0$ contains the following list of sentences

$$(\forall x)(\forall y)[x \in *R] \wedge [y \in *R_n] \Rightarrow [\neg y \in x], \quad n = 0, 1, 2, \dots$$

In words, *there are no internal entities which are elements of the individuals of $*(\hat{R})$.*

(ii) One of the axioms of set theory states that the union of the elements of a set is a set. For the set theory of \hat{R} this means that K_0 contains the following infinite list of sentences.

$$\begin{aligned} (\forall z)[z \in R_n] &\Rightarrow (\exists y)[y \in R_n] \wedge (\forall x)[x \in R_n] \Rightarrow [[x \in y] \\ &\Leftrightarrow (\exists u)[u \in R_n] \wedge [u \in z] \wedge [x \in u]]. \end{aligned}$$

Thus from the F.T. we have the following result: *The union of the elements of an internal entity is an internal entity.*

(iii) The power set axiom of set theory states that for every set there exists a set whose elements are the subsets of this set. Thus K_0 contains the following infinite list of sentences.

$$(\forall x)[x \in R_n] \Rightarrow (\exists y)[y \in R_{n+1}] \wedge (\forall z)[z \in R_n] \Rightarrow [[z \in y] \\ \Leftrightarrow [z \subset x]], \quad n = 1, 2, \dots$$

Then the F.T. implies that *the set of all internal entities which are subsets of an internal entity is an internal entity*.

(iv) Lemma 2.1(v) states that the domain and range of every entity of \hat{R} which is a binary relation is an entity of \hat{R} . This again can be expressed by an infinite list of sentences of K_0 .

$$(\forall b)[b \in B_n] \Rightarrow (\exists z)[z \in R_n] \wedge (\forall x)[x \in R_n] \Rightarrow [[x \in z] \\ \Leftrightarrow (\exists y)[y \in R_n] \wedge [(x, y) \in b]]$$

($n = 3, 4, \dots$), where B_n denotes the entity of all binary relations of rank $\leq n$. The F.T. then implies that *the domain and range of any internal binary relation is internal*.

Another remark which is of importance is that if $b \in \hat{R}$ is a binary relation, then any property which b possesses and which can be expressed by sentences of K_0 also holds for $*b$. For instance, if b is an order relation or function or equivalence relation, then $*b$ is an order relation or function or equivalence relation. If, however, $b \in R$ wellorders its domain, then $*b$ wellorders its domain in the sense that every *nonempty internal* subset of the domain of $*b$ has a first element.

(v) From the axioms of set theory it follows that the image of a set under a binary relation is a set. Thus in \hat{R} the following statement holds. If $b \in \hat{R}$ is a binary relation and $a \in \hat{R}$, then $\{y: (\exists x)(x \in a \wedge (x, y) \in b)\} \in \hat{R}$. We leave it now to the reader to show that this statement can be expressed by sentences of K_0 . The F.T. tells us that the following results holds.

The image of an internal entity under an internal binary relation is internal.

In Theorem 3.6 we have shown that the entities of R^I which are elements of an internal entity are internal, and we remarked that a set of internal entities need not be internal (see Section 5). One of the problems in nonstandard analysis is to decide whether certain sets of internal entities are internal or not. As we shall see in the subsequent sections, one of the methods used to decide such a question involves F.T., by showing that the set in question violates a certain property which it should possess, according to the F.T., if it had been internal. Another useful and helpful result in this respect is the following theorem.

THEOREM 3.10. *Let $V = V(x_1, \dots, x_n)$ be an internal wff with the free variables x_1, \dots, x_n , and let $a \in *(\hat{R})$ be an internal entity. Then the set $\{(x_1, \dots, x_n): (x_1, \dots, x_n) \in a \text{ and } V(x_1, \dots, x_n)\}$ is internal.*

Proof. If V has no quantifiers, that is, $V = V(x_1, \dots, x_n, a_1, \dots, a_p)$, where a_1, \dots, a_p are the constants occurring in V which by hypothesis, denote internal entities. Since a is internal, it follows immediately that the mapping $i \rightarrow E(i) = \{(x_1, \dots, x_n): (x_1, \dots, x_n) \in a(i) \text{ and } V(x_1, \dots, x_n, a_1(i), \dots, a_p(i))\}$ is a mapping of T into R_n for some n , and so determines an internal entity which we shall denote by E . Then it is easy to see that $E = \{(x_1, \dots, x_n): (x_1, \dots, x_n) \in a \text{ and } V\}$. This proves the result for internal wff without quantifiers. For general internal wff we shall use again induction on the number of quantifiers. Thus assume that the theorem holds for all internal wff with $\leq n$ quantifiers. Let $V = (qx_{n+1}) \dots (qx_1)W$ be an internal wff with the free variables y_1, \dots, y_p . There is no loss in generality to assume that $(qx_{n+1}) = (\exists x_{n+1})$ with domain $b \in {}^*(\hat{R})$. Since b is internal it follows from the induction hypothesis that the binary relation

$$B = \{((y_1, \dots, y_p), x_{n+1}): ((y_1, \dots, y_p), x_{n+1}) \in a \times b \text{ and } (qx_n) \dots (qx_1)W(y_1, \dots, y_p, x_{n+1})\}$$

is internal, and so, by Example 3.9(iv), its domain

$$\{(y_1, \dots, y_p): (y_1, \dots, y_p) \in a \text{ and } (\exists x_{n+1})(qx_n) \dots (qx_1)W\}$$

is internal, and the proof is finished.

4. The nonstandard real number system *R . The set *R of individuals of the \mathcal{U} -ultrapower ${}^*(\hat{R})$ of the superstructure \hat{R} , where \mathcal{U} is a δ -incomplete ultrafilter, has according to the F.T. the same properties as R as far as they can be expressed by sentences of K_0 .

Since R is a totally ordered field and since it is easy to see that this can be expressed by sentences of K_0 it follows that *R is a totally ordered field. The imbedding $a \rightarrow {}^*a$ of R into *R imbeds R into a subfield of *R . In order to simplify our notation we shall denote the extensions of the algebraic operation and order when passing from R to *R by the same symbols. Thus $a + b = c$ in *R means in terms of \mathcal{U} that $\{i: a(i) + b(i) = c(i)\} \in \mathcal{U}$, and similarly for subtraction and multiplication. Furthermore, $a \leq b$ in *R means $\{i: a(i) \leq b(i)\} \in \mathcal{U}$. As an illustration the statement that the order relation " \leq " totally orders R can be expressed by the following sentence of K_0

$$(\forall x)(\forall y)[x \in R \wedge y \in R] \Rightarrow [x < y] \vee [x = y] \vee [x > y],$$

and so, as already mentioned above it follows from the F.T. that the extension of the order relation to *R totally orders R .

The unit element $e \in {}^*R$ has the property that for all $0 \neq r \in R$, ${}^*r({}^*r)^{-1} = e$, and so $e = {}^*1$, where 1 denotes the real number one.

The reader will appreciate that we shall simplify our notation further by no longer using the * -notation to denote the **standard individuals** of *R . Thus we

shall from now on identify R with the subfield of the standard numbers of $*R$, and we shall feel free to write $R \subset *R$.

The absolute value $|r|$ of a real number $r \in R$ defined by $|r| = r$ whenever $r > 0$ and $|r| = -r$ whenever $r < 0$ can be considered to be a mapping of R into $R^+ = \{r: r \in R \text{ and } r \geq 0\}$ the set of all nonnegative real numbers. The constant of L denoting this mapping extends by passing from \hat{R} to $*(\hat{R})$ to a mapping $*|\cdot|$ of $*R$ into $*(R^+)$ which according to the F.T. has the property that $*|a| = a$ for all $*R \ni a \geq 0$ and $*|a| = -a$ for all $*R \ni a < 0$. Also in this case we shall drop the $*$ -notation and write $|a|$ to denote the absolute value of a real number $a \in *R$. Similarly, we shall write $\max(a, b)$ and $\min(a, b)$, $a, b \in *R$, for the extensions $*\max(,)$ and $*\min(,)$ of the mappings $\max(r, s)$ and $\min(r, s)$ of $R \times R$ into R respectively.

This liberalization of the notation and some additional notation later on will help a great deal to simplify the mechanics of the subject and can hardly be expected to cause confusion.

Let the constant S denote a subset of R . Then on passing to $*(\hat{R})$, $*S$ denotes a subset of $*R$ which is a standard entity and which by the F.T. has the same properties as S as far as they can be expressed by sentences of K_0 . More precisely the substructure $*(\hat{S})$ of $*(\hat{R})$, where \hat{S} denotes the superstructure defined by S , is an ultrapower nonstandard model of \hat{S} . On the basis of Lemma 3.2(iii) and the present notation, we feel free to write $S \subset *S$. Furthermore, by Lemma 3.2(v), $S = *S$ if and only if S is a finite set.

If the constant N denotes the set of natural numbers of R , that is, $N = \{1, 2, \dots\}$, then the standard entity $*N$ denotes a set of numbers of $*R$ which again has the same properties as N as far as they can be expressed by sentences of K_0 . More precisely, $*(\hat{N})$ is an ultrapower higher order nonstandard model of arithmetic.

From Theorem 3.5 it follows that $*R$ is a proper extension of R , and so, according to a result from algebra to the effect that every Archimedean field is isomorphic to a subfield of R , we conclude that $*R$ is non-Archimedean. But $*R$ has the same properties as R and R is Archimedean. Let us now examine this apparent paradox. The fact that R is Archimedean can be expressed by the following sentence of K_0 :

$$(\forall x)[x \in R] \Rightarrow (\forall n)[n \in N] \Rightarrow [[nx \leq 1] \Leftrightarrow [x \leq 0]],$$

and so, by the F.T., the following statement holds for $*R$.

$$(\forall x)[x \in *R] \Rightarrow (\forall n)[n \in *N] \Rightarrow [[nx \leq 1] \Leftrightarrow [x \leq 0]],$$

that is, with the proper interpretation of the constants, $*R$ is Archimedean with respect to $*N$. It is not Archimedean in the sense of the metalanguage, that is, if $0 < a \in *R$, then there exists a natural number n in the metalanguage such that $a + \dots + a > 1$, n -times +.

Up till now we have only considered some properties of R and their extensions which can be formulated in a lower order language, that is, sentences in which

quantification is over numbers only. Let us now examine a few of the higher order type properties of R . One of the important higher order properties which R possesses and which we have already referred to in the beginning of Section 3 is the so-called **Dedekind completeness property** of R which states that *every nonempty subset of R which is bounded above has a least upper bound*. This statement about R can easily be expressed by a sentence of K_0 which will contain a universal quantifier ranging over subsets of R . Then it follows from the F.T. that $*R$ satisfies a Dedekind completeness property of the following kind.

(4.1) *Every nonempty internal subset of $*R$ which is bounded above has a least upper bound.*

Since $*(\hat{N})$ is a higher order nonstandard model of arithmetic, it follows that under the appropriate interpretation of the F.T. the model $*(\hat{N})$ satisfies all the axioms of Peano. For instance, the principle of induction stating that every nonempty set of natural numbers has a first element, being a higher order property of \hat{N} , has to be interpreted in $*(\hat{N})$ in the following sense.

(4.2) *Every nonempty internal subset of $*N$ has a first element.*

From Theorem 3.5 it also follows that $*N - N \neq \emptyset$. More precisely, we shall now show that there exists a natural number $\omega \in *N$ such that $|r| < \omega$ for all $r \in R$. Indeed, if $\omega(i) = n$ for all $i \in I_n$ ($n = 1, 2, \dots$), where $\{I_n\}$ denotes the partition of I such that $I_n \notin \mathcal{U}$ for all $n = 1, 2, \dots$, then ω is a mapping of I into N with the property that for all $0 < r \in N$ the set $\{i: \omega(i) < r\} \notin \mathcal{U}$, and so $\omega \in *N$ and $|r| < \omega$ for all $r \in R$. This proves on the basis that \mathcal{U} is δ -incomplete that $*N$ contains a number which is larger than any positive real number, that is a number which could be called infinitely large. The reader will find it easy now to appreciate the following definition and facts about $*R$.

DEFINITION 4.3. *A real number $a \in *R$ is called finite whenever there exists a standard real number $0 < r \in R$ such that $|a| < r$. A real number $a \in *R$ which is not finite will be called infinite.*

*A real number $a \in *R$ is called an infinitesimal or infinitely small whenever $|a| < r$ for all $0 < r \in R$.*

The set of all finite real numbers of $*R$ will be denoted by M_0 and the set of all infinitesimals by M_1 .

Observe that $R \subset M_0$, $M_1 \subset M_0$ and $R \cap M_1 = \{0\}$, that is, 0 ("null") being regarded also as an infinitesimal is the *only* standard infinitesimal.

A real number $a \in *R$ is infinite if and only if $|a| > r$ for all $0 < r \in R$. Thus the natural number ω defined above is infinite. Its reciprocal, however, is an infinitesimal. More generally, a real number $0 \neq a \in *R$ is an infinitesimal if and only if its reciprocal $1/a$ is infinite.

The finite natural numbers are determined in the following theorem.

THEOREM 4.4. *A natural number $n \in {}^*N$ is finite if and only if n is a standard natural number. In symbols, ${}^*N \cap M_0 = N$.*

Proof. It is obvious that $N \subset M_0$. If $n \in {}^*N$ is finite, then there exists a standard real number $0 < r \in R$ such that $n < r$. However, K_0 contains the sentence

$$(\forall x)[x \in N] \Rightarrow [x \leq r] \Leftrightarrow [x = 1] \vee [x = 2] \vee \cdots \vee [x = p],$$

where r and p are constants and $p = [r]$ is the integral part of r . Thus by the F.T. we obtain that $n = 1$ or $n = 2$ or \cdots or $n = [r]$, and the proof is complete.

From Theorem 4.4 it follows that the set of all infinitely large natural numbers is given by ${}^*N - N$. It is not uncustomary to denote infinitely large natural numbers by lower case greek letters, such as ω , with or without subscripts.

The mapping $r \rightarrow [r]$ of R_+ into the set $N \cup \{0\}$, where $[r]$ denotes the largest nonnegative integer less than or equal to r , extends on passing from \hat{R} to ${}^*(\hat{R})$ to a mapping ${}^*[\cdot]$ of ${}^*(R_+)$ into ${}^*N \cup \{0\}$. From the F.T. it follows that for all $0 \leq a \in {}^*R$, ${}^*[a]$ is the largest nonnegative integer $\leq a$. Also in this case we shall drop the * -notation and simply write $[a]$ for the integral part of a .

We shall now turn to a discussion of the properties of the finite numbers of *R .

It is easy to see that M_0 is a subring of *R , and in fact is an integral domain, that is, M_0 has no divisors of zero. The set of infinitesimals constitutes a subring of M_0 with the property that if $h \in M_1$ and $a \in M_0$, then $ah \in M_1$, that is M_1 is an ideal in M_0 . In fact, it is easy to see that M_1 is a maximal ideal. Indeed, observe that if $a \in M_0$ and $a \notin M_1$, then there exist positive real numbers $r_1, r_2 \in R$ such that $0 < r_1 < |a| < r_2$, and so $1/a \in M_0$ shows that any ideal which properly contains M_1 must contain the unit element 1 of M_0 and so is all of M_0 .

If $a, b \in {}^*R$ and $a - b$ is infinitesimal, then we shall say that b is **infinitely close** to a and we write $a = {}_1b$.

Consider the quotient ring M_0/M_1 . Then since M_1 is a maximal ideal in M_0 , the quotient ring M_0/M_1 is a field. We claim it is isomorphic to the field of standard real numbers. The precise result and details are the subject of the following important theorem.

THEOREM 4.5. *The quotient ring M_0/M_1 is order isomorphic to the field R of the standard real numbers.*

Proof. First observe that if A is an equivalence class in M_0 modulo M_1 , then A cannot contain two different standard real numbers r_1 and r_2 . Indeed, in that case $|r_1 - r_2| = {}_10$, and so $r_1 \neq r_2$ implies by Definition 4.4 that $|r_1 - r_2| < |r_1 - r_2|$ and a contradiction is obtained. This shows that R is a subfield of M_0/M_1 . To complete the proof we have to show that to every $a \in M_0$ there corresponds a standard real number r , which is then unique, such that $a - r = {}_10$. To this end, observe that if $a \in M_0$, then the sets $D = \{r: r \in R \text{ and } r \leq a\}$ and $D' = R - D$ define a Dedekind cut (D, D') in R . Let $r \in R$ be the real number in R which deter-

mines the same cut (D, D') . Then we shall show that $a = {}_1r$. If not, then by Definition 4.4 there exists a positive real number $0 < \varepsilon \in R$ such that $|a - r| \geq \varepsilon$. If $a > r$, then $|a - r| \geq \varepsilon$ implies that $r + \varepsilon/2 < a$, and contradicts the fact that a and r determine the same cut. Similarly, if $r > a$, then $r - \varepsilon/2 > a$ gives rise to the same contradiction. Thus M_0/M_1 is order isomorphic to R and the proof is finished.

The unique ring and order isomorphism of M_0 onto R with kernel M_1 plays a very important role in the theory of infinitely small and infinitely large numbers. We shall firmly establish it in the following definition.

DEFINITION 4.6. *The ring and order homomorphism of M_0 onto R_0 with kernel M_1 will be called the standard part homomorphism and will be denoted by st .*

In the next theorem, we shall summarize the basic properties of the homomorphism st for later reference.

THEOREM 4.7. (i) $st(a + b) = st(a) + st(b)$, $st(ab) = st(a)st(b)$ and $st(a - b) = st(a) - st(b)$ for all $a, b \in M_0$.

(ii) If $a, b \in M_0$, then $a \leq b$ implies $st(a) \leq st(b)$.

(iii) $st(|a|) = |st(a)|$, $st(\max(a, b)) = \max(st(a), st(b))$ and $st(\min(a, b)) = \min(st(a), st(b))$ for all $a, b \in M_0$.

(iv) $st(a) = 0$ if and only if $a \in M_1$.

(v) For all standard $r \in R$ we have $st(r) = r$.

(vi) If $a \in M_0$ and $st(a) \geq 0$, then $|a| = {}_1st(a)$.

(vii) For all $a, b \in M_0$ we have $a = {}_1b$ if and only if $st(a) = st(b)$.

It is now customary to call the equivalence classes of M_0 with respect to M_1 the **monads** of the standard numbers determined by them. The monads are denoted by $\mu(r)$, $r \in R$. Thus, in particular, $\mu(0) = M_1$.

We shall conclude this section with a number of remarks which are of interest in themselves.

REMARKS. (i) (*The standard part operation defined as a limit*). The standard part operation " st " can also be defined as follows. If $a \in M_0$, then, by Definition 4.4 and Definition 3.1, there is a set $U \in \mathcal{U}$ and a positive standard real number $0 < r \in R$ such that $i \in U$ implies $|a(i)| < r$. Hence, the image of the ultrafilter \mathcal{U} under the mapping $i \rightarrow a(i)$ of I into R is a basis of a bounded ultrafilter of subsets of R , and so, by the local compactness of R , it converges to a unique real number r . A simple observation shows that $r = st(a)$. Thus, $st(a) = \lim_{\mathcal{U}} a$ for all $a \in M_0$.

(ii) (*A nonstandard construction of the real number systems*). The proof of Theorem 4.5 suggests immediately the following alternative construction of the real number system. Let the constant Q of L denote the field of rational numbers. Then ${}^*(\hat{Q})$ is a higher order nonstandard model of the superstructure Q . Thus the set of individuals ${}^*Q \subset {}^*R$ is a subfield of *R which has the same properties as Q as far as they can be expressed by sentences of K_0 . From Theorem 3.5 we know that

$*Q \neq Q$, and in fact $*Q$ contains an element which is larger than any standard real number. It is an easy and interesting exercise for the reader to transform the properties of Q to $*Q$. We shall show here only that $*Q$ can be used to define the real number system. To this end, we single out the rationals of $*Q$ which are finite, that is, $q \in *Q$ is finite whenever $|q| < \text{some positive standard rational number}$. The set of all finite rationals will be denoted by Q_0 . Observe that $Q_0 = *Q \cap M_0$. A rational $q \in *Q$ is called infinitesimal whenever $|q|$ is smaller than all positive standard rationals. The set of all infinitely small rationals will be denoted by Q_1 . Thus $Q_1 = *Q \cap M_1$. Then it is easy to see that Q_1 is a maximal ideal in the integral domain Q_0 . Thus the quotient ring Q_0/Q_1 is ring and order isomorphic to a field. The proof of Theorem 4.5 shows us, however, that this field is isomorphic to the field of Dedekind cuts of Q , and so, by definition, Q_0/Q_1 is isomorphic to the real number system.

(iii) (*The nonstandard complex number system*). Within the framework of axiomatic set theory the complex number system C may be regarded as a subtheory of the theory of the superstructure $R \times R$ determined by $R \times R$. The algebraic operations of addition and multiplication are denoted by constants which correspond to certain six-place relations; and so $*(R \times R)$ may be looked upon as a higher order non-standard model of the complex number system.

It is advisable also in this case to employ the familiar notation $z = x + iy$ for complex numbers, where now $x, y \in *R$ and $i^2 = -1$. The set $*C = *R \times *R$ of the extended complex number system has of course the same properties as C , and so is, in particular, a field. If $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$, then also in $*C$ we have

$$z_1 + z_2 = x_1 + x_2 + i(y_1 + y_2) \text{ and } z_1 z_2 = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1).$$

Furthermore, $z = x + iy$, then x is called the real part of z and y is called the imaginary part of z . A complex number $z = x + iy$ is finite whenever x and y are finite, otherwise it is infinite. If x and y are both infinitely small, then $z = x + iy$ is called an infinitely small complex number. From Theorem 4.6 it follows that every finite complex number is infinitely close to a unique standard complex number. For further details concerning nonstandard complex function theory we refer the reader to [8] and [10].

5. Definitions and properties of some external entities. We pointed out that the converse of Theorem 3.6 need not hold, that is, a set of internal entities need not be internal. In the preceding section we introduced a number of sets of individuals, namely, the set of all infinitely large natural numbers $*N - N$, the set of finite numbers M_0 , the set of infinitesimals M_1 , and the monads $\mu(r)$, $r \in R$. It is now natural to ask the question whether these sets are internal or not? To decide this we shall use the following procedure. We assume the set in question is internal and

then show that it violates a property which it should have possessed on the basis of the assumption that it is internal and the F.T. The details are contained in the following theorem.

THEOREM 5.1. *The nonempty sets $*N - N$, M_0 , M_1 , $\mu(r)$ ($r \in R$), and the set of infinitely large real numbers $*R_\infty = *R - M_0$ are all external.*

Proof. Assume that $*N - N$ is internal. Then since $*N - N \neq \emptyset$ (Theorem 3.5) we have by (4.2) that $*N - N$ has a first element, say, ω_0 . But the set of infinitely large natural numbers does not have a first element. Indeed, if $\omega \in *N - N$, then $k + 1 < \omega$ for all $k \in N$ implies that $\omega - 1 \in *N - N$, and so $\omega_0 - 1 < \omega_0$ shows that $*N - N$ has no first element. Thus $*N - N$ is external.

Assume that the set M_1 is internal. Since $M_1 \neq \emptyset$ and $h \in M_1$ implies $|h| < 1$ it follows from (4.1) that M_1 has a least upperbound, say, a_0 . From $0 \in M$, it follows that $a_0 \geq 0$. Furthermore, $a_0 \notin M_1$ since M_1 contains elements other than 0. But then $a_0/2$ is also a least upper bound of M_1 and a contradiction is obtained, and so M_1 is external.

Similarly on the basis of (4.1) we can show that M_0 is external. We leave it to the reader as an exercise.

If $*R_\infty = *R - M_0$ is internal, then also $M_0 = *R - *R_\infty$ is internal, and a contradiction is obtained. Thus $*R_\infty$ is external.

Since the translation mappings of $*R$ are internal (check this) it follows immediately from $\mu(0)$ is external that $\mu(r) = \mu(0) + r$ ($r \in R$) is external. This completes the proof.

REMARKS. (i) If $D \subset M_1$ is internal and nonempty, then according to (4.1) it has a least upper bound. The above proof shows that this least upper bound is an infinitesimal. Similarly, the least upper bound of a nonempty internal set of finite numbers is finite. The greatest lower bound of a nonempty internal set of infinite numbers is of course infinite.

(ii) The standard part operation is a mapping of M_0 onto R . It is, however, *not an internal mapping*. Indeed, if it were internal, then according to Example 3.9(v) its domain M_0 would have to be internal which contradicts the preceding theorem, and we conclude that the standard part operation is an external operation.

Let $A \in \hat{R}$ be infinite. Then according to Theorem 3.5 the set of all the nonstandard entities of $*A$ is not empty. More precisely, we have the following result.

THEOREM 5.2. *If $A \in \hat{R}$, then the set $*A - \{a : a \in A\}$ of all the nonstandard elements of $*A$ is either empty or external, and in the latter case the set $\{a : a \in A\}$ is also external.*

Proof. If $A \in \hat{R}$, then $*A - \{a : a \in A\} = \emptyset$ if and only if A is finite. (Theorem 3.5 and Lemma 3.2(v)). Assume therefore that A is infinite. Then there is a one-to-one mapping f of a subset of A onto the set $N = \{1, 2, \dots\}$. If $B = *A - \{a : a \in A\}$ is

internal, then $B \cap \text{dom}(*f)$ is internal also (Theorem 3.10). Hence, by Example 3.9(v), we have that $*N - N = *f(B \cap \text{dom} *f)$ is internal which contradicts Theorem 5.1 and the proof is finished.

Although the preceding theorem shows that the set of nonstandard elements of the extension of an infinite set of \hat{R} is external there are plenty of internal sets whose elements are all internal entities which are not standard. Indeed, if $\omega \in *N - N$, then the set $\{\omega\}$ is internal but its element is not a standard entity. More generally any finite set of internal entities which are not standard is internal. This statement can be generalized as follows. We begin with a definition.

DEFINITION 5.3. *A set D of internal entities of $*(\hat{R})$ is called $*\text{-finite}$ whenever there exists a natural number $\omega \in *N - N$ and an internal one-to-one mapping of D onto the internal set $\{1, 2, \dots, \omega\}$. In that case, we shall say that the internal cardinal of D is ω or shortly that D has ω -elements.*

If D is $*\text{-finite}$, then it is clear that its external cardinal is at least as big as \aleph_0 . Concerning $*\text{-finite}$ sets we have the following result.

THEOREM 5.4. *Every $*\text{-finite}$ set of internal entities is internal. A $*\text{-finite}$ set of real numbers has a largest and a smallest element.*

Proof. Since, by Example 3.9(iv), the domain of an internal function is internal it follows immediately from Definition 5.3 that a $*\text{-finite}$ set is internal.

If D is a $*\text{-finite}$ set of real numbers, then from the sentence of K_0 stating that every finite set of real numbers of R has a largest and a smallest element it follows from the F.T. that every $*\text{-finite}$ set of real numbers in $*R$ has a largest and a smallest element. This completes the proof.

REMARK. If the internal set D is $*\text{-finite}$, then it must contain at least one internal entity which is not standard, and so at least externally infinitely many of those. This can be shown as follows. If the entities of D are all standard, then there exists a standard set $A \in \hat{R}$ such that $D = \{ *a : a \in A \}$ (use first part of Theorem 3.6). Since the cardinal of A is infinite it follows from Theorem 5.2 that the set $D = \{ *a : a \in A \}$ is external, and so a contradiction is obtained.

6. The theory of limits. As a first example and also for later reference we shall illustrate what kind of effect the theory of infinitely small and infinitely large numbers has on the theory of limits.

We recall that a (standard) sequence $\{s_n : n = 1, 2, \dots\}$ can be regarded as a mapping of N into R , and so being a subset of $N \times R$ it is an entity of \hat{R} which we shall denote for obvious reasons by s . On passing from \hat{R} to $*(\hat{R})$ the entity s extends to an entity $*s$ which according to the F.T. and Lemma 3.2(vii) is a mapping of $*N$ into $*R$. Furthermore, for all finite $n \in N$ we have $*s_n = s_n$ as follows from the fact that $*(\text{ran } s) = \text{ran } *s$ and the convention of dropping the $*$ -notation for indi-

viduals. The standard sequence $*s$ in $*(\hat{R})$ has the same properties as the sequence s as far as they can be expressed by sentences of K_0 . With this fundamental principle in mind we shall now prove the following theorems.

THEOREM 6.1. *A sequence $\{s_n; n = 1, 2, \dots\}$ in R is bounded if and only if $*s_\omega$ is finite for all infinitely large natural numbers $\omega \in *N - N$.*

Proof. This follows immediately from the remark following Theorem 5.1 to the effect that the least upper bound of an internal set of finite numbers is finite. Hence, if $(\text{ran } *s) \subset M_0$, then $|*s_n| \leq a$ for all $n \in *N$ and some $a \in M_0$, that is, $|s_n| \leq st(a)$ for all $n \in N$, and the proof is finished.

In the classical sense a sequence $\{s_n; n = 1, 2, \dots\}$ is said to be convergent with limit s if and only if

$$(*) \quad (\forall \varepsilon)[0 < \varepsilon \in R] \Rightarrow (\exists x)[x \in N] \wedge (\forall y)[y \in N \wedge x \leq y] \Rightarrow [|s_y - s| < \varepsilon].$$

In nonstandard analysis this is expressed in a more intuitive fashion as follows.

THEOREM 6.2. *Let $\{s_n; n = 1, 2, \dots\}$ be a sequence of numbers of R , and let $s \in R$. Then $\lim_{n \rightarrow \infty} s_n = s$ if and only if $*s_\omega = {}_1s$ for all $\omega \in *N - N$.*

Proof. Assume first that $\lim_{n \rightarrow \infty} s_n = s$. Then from the sentence $(*)$ of K_0 the following is a sentence of K_0 .

$$(\forall x)[x \in N \wedge x > n] \Rightarrow |s_x - s| < \varepsilon, \quad \text{where } \varepsilon > 0 \text{ and } n \in N$$

are constants. Thus the following $*L$ -sentence holds.

$$(\forall x)[x \in *N \wedge x > n] \Rightarrow |*s_x - s| < \varepsilon.$$

In particular, for all $\omega \in *N - N$ we have that $|*s_\omega - s| < \varepsilon$. The latter statement holds, however, for all $\varepsilon > 0$, that is, $*s_\omega = {}_1s$ for all $\omega \in *N - N$.

In order to see that the condition is sufficient we observe that if ε is a constant denoting a positive number of R , the following sentence holds in $*(\hat{R})$.

$$(\exists y)[y \in *N] \wedge (\forall x)[x \in *N \wedge y < x] \Rightarrow |*s_x - s| < \varepsilon.$$

Indeed, we need to take for y only an infinitely large natural number. Observe now that this sentence is the $*$ -transform of the sentence

$$(\exists y)[y \in N] \vee (\forall x)[x \in N \wedge y < x] \Rightarrow |s_x - s| < \varepsilon,$$

and so by F.T. holds in \hat{R} . This means that there is an index $n_0 \in N$ such that $|s_n - s| < \varepsilon$ for all $n > n_0$. Since this holds for all $\varepsilon > 0$ we obtain that $\lim_{n \rightarrow \infty} s_n = s$ and the proof is finished.

The condition $*s_\omega = {}_1s$ for all $\omega \in *N - N$ is equivalent to $st(*s_\omega) = s$ for all $\omega \in *N - N$.

Theorem 6.2 also tells us immediately that if the limit exists it is unique. Furthermore, Theorem 6.1 shows that every convergent sequence is bounded.

EXAMPLES 6.3. (i) If one wishes to show that $\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1$, then set $s_n = \sqrt[n]{n} - 1$ ($n = 1, 2, \dots$) and observe that

$$n = (1 + s_n)^n = \sum_{k=0}^n \binom{n}{k} s_n^k \geq \binom{n}{2} s_n^2 \quad \text{for all } n = 1, 2, \dots$$

Hence, $0 \leq s_n \leq \sqrt{2/(n-1)}$ for all $n > 1$, and so, also $0 \leq {}^*s_m \leq \sqrt{2/(m-1)}$ for all $1 < m \in {}^*N$. In particular if $\omega \in {}^*N - N$ is infinitely large, then $0 \leq {}^*s_\omega \leq \sqrt{2/(\omega-1)}$ and $\sqrt{2/(\omega-1)} \in M_1$ implies that ${}^*s_\omega = {}_10$, and so, by Theorem 6.2, $\lim_{n \rightarrow \infty} s_n = 0$, and the proof is finished.

(ii) (*Algebra of limits*). The usual rules for calculating with limits are now easily obtained. For, if $\lim s_n = s$ and $\lim t_n = t$, then ${}^*(s+t)_\omega = {}^*s_\omega + {}^*t_\omega = {}_1s + t$ for all $\omega \in {}^*N - N$ and so $\lim_{n \rightarrow \infty} (s_n + t_n) = s + t$. Similarly, $st({}^*(st)_\omega) = st({}^*s_\omega {}^*t_\omega) = st({}^*s_\omega)st({}^*t_\omega) = st$ for all $\omega \in {}^*N - N$ shows that $\lim_{n \rightarrow \infty} s_n t_n = st$. In the same way one shows that if $t \neq 0$, then $\lim_{n \rightarrow \infty} s_n/t_n = s/t$.

(iii) It is well known that if $\lim_{n \rightarrow \infty} s_n = s$, then

$$\lim_{n \rightarrow \infty} \frac{s_1 + \dots + s_n}{n} = s.$$

The proof of this result in nonstandard analysis reads as follows. From $\lim_{n \rightarrow \infty} s_n = s$ it follows first of all that for some $0 < r \in R$, $|{}^*s_n - s| < r$ for all $n \in {}^*N$ (Theorem 6.1) and ${}^*s_n - s = {}_10$ for all $n \in {}^*N - N$. Now let $\omega \in {}^*N - N$ and let $\omega_0 = [\sqrt{\omega}]$. Then the following simple estimation gives the required result:

$$\begin{aligned} \left| \frac{{}^*s_1 + \dots + {}^*s_\omega}{\omega} - s \right| &\leq \frac{|{}^*s_1 - s| + \dots + |{}^*s_{\omega_0} - s|}{\omega_0} \cdot \frac{1}{\sqrt{\omega}} \\ &\quad + \frac{|{}^*s_{\omega_0+1} - s| + \dots + |{}^*s_\omega - s|}{\omega} \\ &\leq \frac{r}{\sqrt{\omega}} + \frac{(\omega - \omega_0)}{\omega} \cdot \max(|{}^*s_n - s| : \omega_0 < n \leq \omega) = {}_10, \end{aligned}$$

by Theorem 5.4.

Cauchy's criterion for convergence in analysis takes on the following form.

THEOREM 6.4. *A sequence $\{s_n : n = 1, 2, \dots\}$ of real numbers of R is convergent if and only if ${}^*s_\omega = {}_1{}^*s_{\omega'}$ for all $\omega, \omega' \in {}^*N - N$.*

Proof. From Cauchy's criterion $|s_n - s_m| < \varepsilon$ for all n, m sufficiently large it follows as in the proof of Theorem 6.2 that the condition is necessary. In order to

prove that the condition is sufficient we have only to show in view of Theorem 6.2 that $*s_\omega$ is finite for all $\omega \in *N - N$. To this end, assume that there exists an infinitely large natural number $\omega_0 \in *N - N$ such that $*s_{\omega_0}$ is infinite. We define now the following set $A = \{n: *N \text{ and } |*s_{\omega_0} - *s_n| < 1\}$ of natural numbers. From Theorem 3.10 it follows that A is internal. Furthermore, by hypothesis $*N - N \subset A$. If $n \in N$ is finite, then $|*s_{\omega_0}| \leq |*s_{\omega_0} - *s_n| + |*s_n| \in M_0$ shows that $n \notin A$, and so $A = *N - N$. Contradicting the fact that $*N - N$ is not internal (Theorem 5.1), and so $*s_\omega$ is finite for all $\omega \in *N - N$, and the proof is finished.

REMARK. The above proof shows also that an infinite sequence $\{s_n: n = 1, 2, \dots\}$ is bounded if and only if $*s_\omega - *s_{\omega'}$ is finite for all $\omega, \omega' \in *N - N$.

The following result of A. Robinson (see [9]) concerning internal sequences will be used in Section 9.

THEOREM 6.5. *Let $\{a_n: n \in *N\}$ be an internal sequence of real numbers such that a_n is infinitely small for all finite $n \in N$. Then there exists an infinitely large natural number $\omega \in *N - N$ such that $a_n = {}_1 0$ for all $n \leq \omega$.*

Proof. Consider the internal sequence $\{na_n: n \in *N\}$ and let $A = \{n: n \in *N \text{ and } \forall k [k \in *N \wedge k \leq n] \Rightarrow k | a_k| \leq 1\}$. Then, by Theorem 3.10, A is internal. Since the hypothesis $a_n = {}_1 0$ for all finite $n \in N$ implies $na_n = {}_1 0$ for all finite $n \in N$ it follows that $N \subset A$. Since, by Theorem 5.2, the set N is external and since A is internal, $A - N \neq \emptyset$. Hence, there exists an infinitely large natural number $\omega \in A$. Then for all infinitely large $n \leq \omega$ the condition $n | a_n| \leq 1$ implies that $0 \leq |a_n| \leq 1/n = {}_1 0$, and the proof is finished.

7. Sequences that are asymptotically linear. A standard sequence of real numbers $\{s_n: n = 1, 2, \dots\}$ is called **asymptotically linear** whenever there exists a real constant $\sigma \in R$ such that $s_n = n\sigma + o(n)$, $n \in N$.

A now classical result of Pólya and Szegő states if a sequence $\{s_n: n = 1, 2, \dots\}$ is almost additive, that is, there exists a constant s such that $|s_{n+m} - s_n - s_m| \leq s$ for all $n, m = 1, 2, \dots$, then $\{s_n\}$ is asymptotically linear.

As another illustration of the use of infinitely small and infinitely large numbers we shall prove here in a nonstandard fashion the following slightly more general result.

THEOREM 7.1. *Let $\{s_n: n = 1, 2, \dots\}$ be a standard sequence of real numbers for which there exist constants p, s such that $0 < p < 1$ and $|s_{n+m} - s_n - s_m| \leq s(n^p + m^p)$ for all $n, m = 1, 2, \dots$. Then there exists a constant $\sigma \in R$ such that $|s_n - n\sigma| \leq sn^p/(1 - 2^{p-1})$ for all $n = 1, 2, \dots$. In particular, $\{s_n\}$ is asymptotically linear.*

Proof. From the hypothesis it follows immediately that for all $k = 1, 2, \dots$ and for all $n = 1, 2, \dots$ we have

$$(7.2) \quad \left| \frac{s_{2^k n}}{2^k} - s_n \right| \leq sn^p \frac{1 - 2^{(p-1)k}}{1 - 2^{p-1}}.$$

Then it follows from the F.T. that, by passing to $*(\hat{R})$, (7.2) holds for all $k, n \in *N$. In particular, if $k = \omega \in *N - N$ is infinitely large, then

$$(7.3) \quad \left| \frac{*s_{2^\omega n}}{2^\omega} - *s_n \right| \leq sn^p \frac{1 - 2^{(p-1)\omega}}{1 - 2^{p-1}} \quad \text{for all } n \in *N.$$

Since $0 < p < 1$, and ω is infinitely large, $2^{(p-1)\omega}$ is infinitely small, and so for all finite n we see that $*s_{2^\omega n}/2^\omega$ is finite. Let

$$a_n = \frac{*s_{2^\omega n}}{2^\omega}, \quad n \in *N.$$

Then the internal sequence $\{a_n : n \in *N\}$ satisfies, by hypothesis, the condition that $|a_{n+m} - a_n - a_m| \leq s 2^{(p-1)\omega}(n^p + m^p)$ for all $n, m \in *N$. Since a_n is finite for all finite $n \in N$ we obtain by setting $t_n = st(a_n)$, $n \in N$, that $|t_{n+m} - t_n - t_m| = 0$, that is, $t_n = nt_1 = n\sigma$, $n = 1, 2, \dots$. Finally, if we take standard parts in (7.3) keeping n finite we obtain that $|s_n - n\sigma| \leq sn^p/(1 - 2^{p-1})$, and the proof is finished.

8. Continuity and differentiability. Let f be a real-valued function of a real variable which is defined on an open interval $a < x < b$ of R . On passing to $*(\hat{R})$ the function f extends to a function $*f$ whose domain of definition is the open interval $a < x < b$, $x \in *R$ and with values in $*R$. Furthermore, we have to keep in mind that the F.T. implies that $*f$ satisfies in $*(\hat{R})$ all the properties of f as far as they can be expressed by sentences of K_0 .

For instance, if for some $a < x_0 < b$, $\lim_{x \rightarrow x_0} f(x) = l$ holds, then the following sentence belongs to K_0 .

$$\begin{aligned} (\forall \varepsilon)[0 < \varepsilon \in R] &\Rightarrow (\exists \delta)[0 < \delta \in R] \wedge (\forall x)[x \in R \wedge 0 < |x - x_0| < \delta] \\ &\Rightarrow [|f(x) - l| < \varepsilon]. \end{aligned}$$

Using the same methods as in the proof of Theorem 6.1 we obtain immediately the following result.

THEOREM 8.1. $\lim_{x \rightarrow x_0} f(x) = l$ if and only if $*f(x_0 + h) = {}_1 l$ for all $0 \neq h \in M_1$. In particular, f is continuous at x_0 if and only if $*f(x_0 + h) = {}_1 f(x_0)$ for all $h \in M_1$, that is, equivalently, $st(*f(a)) = f(st(a))$ for all $a \in *R$ such that $st(a) = x_0$.

The derivative of f at x_0 exists if and only if

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

exists. Thus, by Theorem 8.1, f is differentiable at x_0 if and only if there exists a con-

stant $l \in R$ such that

$$\frac{{}^*f(x_0 + h) - {}^*f(x_0)}{h} =_1 l$$

for all $0 \neq h \in M_1$. As we might have expected the derivative of a differentiable function is the standard part of the quotient of infinitesimals

$$\frac{\Delta f}{\Delta x} = \frac{{}^*f(x + \Delta x) - f(x)}{\Delta x},$$

where $\Delta x \neq 0$ denotes an infinitesimal.

If f is differentiable at x_0 , then f is continuous at x_0 . Indeed, from ${}^*f(x_0 + h) - f(x_0) =_1 h f'(x_0)$ for all $0 \neq h \in M_1$ it follows, using $hl =_1 0$, that ${}^*f(x_0 + h) - f(x_0) =_1 0$ for all $h \in M_1$.

A real function f defined on an arbitrary interval is uniformly continuous whenever for every $0 < \varepsilon \in R$ there exists a constant $0 < \delta \in R$ such that $|f(x) - f(y)| < \varepsilon$ for all $x, y \in \text{dom } f$ and $|x - y| < \delta$. In passing to ${}^*(\hat{R})$ we obtain immediately the following criterion for uniform continuity.

THEOREM 8.2. *Let f be a real function of a real variable. Then f is uniformly continuous if and only if ${}^*f(a) =_1 {}^*f(b)$ for all $a, b \in \text{dom } {}^*f$ and $a =_1 b$.*

From the above results the following famous theorem of Heine can now be obtained immediately.

THEOREM 8.3. (Heine). *Let f be a real function of a real variable defined on the bounded and closed interval $x_1 \leq x \leq x_2$, $x_1, x_2 \in R$. If f is continuous, then f is uniformly continuous.*

Proof. Let $a, b \in {}^*R$ satisfy $x_1 \leq a, b \leq x_2$ and $a =_1 b$. Then $a, b \in M_0$ and $x = st(a) = st(b)$ satisfies $x_1 \leq x \leq x_2$. Since f is continuous we have, by Theorem 8.1, that ${}^*f(a) =_1 f(x) =_1 {}^*f(b)$, and so ${}^*f(a) =_1 {}^*f(b)$, that is, by Theorem 8.2, f is uniformly continuous, and the proof is finished.

For a more detailed account of the theory of real functions of a real variable in non-standard analysis we refer the reader to [3] and [8].

9. Euler's product for the sine function. On passing from \hat{R} to ${}^*(\hat{R})$, the elementary functions of the calculus such as the functions $\log x$, e^x , $\sin x$, $\cos x$, and so on, extend to functions defined in *R and which have the same properties as their standard counterpart as far as they can be expressed by sentences of K_0 . In order to simplify the notation we shall not use the $*$ -notation to denote the extensions of the elementary functions. Thus, for instance, in place of writing ${}^*(\sin)(x)$, $x \in {}^*R$, we simply write $\sin x$, $x \in {}^*R$. For a discussion of the elementary functions of ${}^*(\hat{R})$ we refer the reader to [3].

One of the many beautiful formulas which were discovered by Euler is the so-called product formula for the sine-function. By this we mean the following formula.

$$(9.1) \quad \sin z = z \prod_{k=1}^{\infty} \left(1 - \frac{z^2}{k^2\pi^2}\right), \quad z \text{ is complex.}$$

Nowadays this representation for the sine function belongs to that part of function theory that studies the behavior of entire functions whenever its zeros are given. There one learns that the quotient of the functions on the left and right-hand side of (9.1) is a function of the form e^f , where f is entire. The whole problem is then to determine f , and, in fact, to show that $f = 0$ in the case of the sine function. There are many proofs known for this result. Some of the proofs are even elementary. But all of these proofs are somewhat artificial in the sense that they rely on some analytical trick. It is therefore not without interest to examine how Euler proved his formula. As far as the author knows, Euler's original proof is contained in his book *Introductio ad Analysin Infinitorum* which appeared in 1748. It runs as follows. The mathematical expressions such as "infinitely large" and "infinitely close" which occur in it are Euler's and not the author's.

For infinitely large values of n we have

$$(9.2) \quad 2 \sinh x = \left(1 + \frac{x}{n}\right)^n - \left(1 - \frac{x}{n}\right)^n.$$

We are now going to factorize the polynomial occurring on the right-hand side of (9.2), by observing that $a^n - b^n = (a-b)(a-\varepsilon_1 b) \cdots (a-\varepsilon_{n-1} b)$, where $1, \varepsilon_1, \dots, \varepsilon_{n-1}$ are the n th roots of unity. Now combine the pairs of complex conjugate roots to obtain the real quadratic polynomials

$$\left(a - b \exp\left(\frac{2k\pi i}{n}\right)\right) \left(a - b \exp\left(-\frac{2k\pi i}{n}\right)\right) = a^2 + b^2 - 2ab \cos \frac{2k\pi}{n},$$

and so, since $a^2 + b^2 = 2 + (2x^2/n^2)$ and $2ab = 2 - (2x^2/n^2)$, we obtain

$$2 \left(1 - \cos \frac{2k\pi}{n}\right) + 2 \left(1 + \cos \frac{2k\pi}{n}\right) \frac{x^2}{n^2} = 4 \sin^2 \frac{k\pi}{n} \left(1 + \frac{x^2}{n^2 \tan^2(k\pi/n)}\right).$$

It follows that the polynomial is divisible by x and for all values of $k = 1, 2, \dots$, by $1 + \{x^2/n^2 \tan^2(k\pi/n)\}$. Since n is infinitely large this factor is infinitely close to $1 + (x^2/k^2\pi^2)$. Furthermore, it is easy to see that the coefficient of x is equal to 2, and so we obtain that

$$(9.3) \quad \sinh x = x \prod_{k=1}^{\infty} \left(1 + \frac{x^2}{k^2\pi^2}\right).$$

Finally, by applying it for $x = iz$, the required formula is obtained.

The reader who has read this far will agree with the author that Euler's proof is a typical example of the way infinitely large and infinitely small numbers were used with great success in the early stages of the development of the calculus. It is, however, no wonder that the inability to give the theory of infinitely large and infinitely small numbers a firm foundation led to the unacceptability of such proofs. Of course, it is no problem at all with the methods of nonstandard analysis to make Euler's proof precise.

From Theorem 6.1 it follows that for all standard $x \in R$ and for all infinitely large natural numbers $\omega \in {}^*N - N$ we have

$$(9.4) \quad 2 \sinh x = {}_1 \left(1 + \frac{x}{\omega} \right)^\omega - \left(1 - \frac{x}{\omega} \right)^\omega.$$

Factorizing the polynomial as before leads to the formula.

$$(9.5) \quad \left(1 + \frac{a}{m} \right)^m - \left(1 - \frac{a}{m} \right)^m = \frac{4^{m/2}}{m} \left(\sum_{k=1}^{[(m-1)/2]} \sin^2 \frac{k\pi}{m} \right) a \sum_{k=1}^{[(m-1)/2]} \left(1 + \frac{a^2}{m^2 \tan^2 \frac{k\pi}{m}} \right),$$

for all $a \in {}^*R$ and for all $m \in {}^*N$, and where $[(m-1)/2]$ as in Section 4 denotes the largest natural number $\leq (m-1)/2$. Dividing by $a \neq 0$ and letting $a = 0$ shows that

$$(9.6) \quad \frac{4^{m/2}}{m} \prod_{k=1}^{[(m-1)/2]} \sin^2 \frac{k\pi}{m} = 2 \quad \text{for all } m \in {}^*N.$$

Thus we obtain finally that

$$(9.7) \quad \left(1 + \frac{x}{\omega} \right)^\omega - \left(1 - \frac{x}{\omega} \right)^\omega = 2x \prod_{k=1}^{[(\omega-1)/2]} \left(1 + \frac{x^2}{\omega^2 \tan^2(k\pi/\omega)} \right),$$

or all $x \in R$ and for all $\omega \in {}^*N - N$.

We shall now prove the following lemma.

LEMMA 9.8. *If $x \in R$ is standard, then for all infinitely large $\omega \in {}^*N - N$ we have*

$$st \left(\prod_{k=1}^{[(\omega-1)/2]} \left(1 + \frac{x^2}{\omega^2 \tan^2(k\pi/\omega)} \right) \right) = \prod_{k=1}^{\infty} \left(1 + \frac{x^2}{k^2 \pi^2} \right).$$

Proof. Since for all $x \in R$, the infinite product $\prod_{k=1}^{\infty} (1 + x^2/k^2 \pi^2)$ is convergent it follows from Theorem 6.1 that

$$(9.9) \quad \prod_{k=1}^{\infty} \left(1 + \frac{x^2}{k^2 \pi^2} \right) = {}_1 \prod_{k=1}^{[(\omega-1)/2]} \left(1 + \frac{x^2}{k^2 \pi^2} \right)$$

for all $x \in R$ and for all $\omega \in {}^*N - N$.

Since $\omega^2 \tan^2(k\pi/\omega) \geq k^2\pi^2$ for all $1 \leq k \leq [(\omega-1)/2]$ we obtain that

$$\sum_{k=1}^{[(\omega-1)/2]} \left(\log \left(1 + \frac{x^2}{k^2\pi^2} \right) - \log \left(1 + \frac{x^2}{\omega^2 \tan^2(k\pi/\omega)} \right) \right) \geq 0, \text{ for all } x \in R.$$

From Theorem 3.10 it follows that the following sequence is internal

$$(9.10) \quad \eta_n = \sum_{k=1}^n \left(\log \left(1 + \frac{x^2}{k^2\pi^2} \right) - \log \left(1 + \frac{x^2}{\omega^2 \tan^2(k\pi/\omega)} \right) \right), \quad n \in {}^*N \text{ and } x \in R.$$

If n is finite, then, by Theorem 8.1, the continuity of the log-function and $n/\omega = {}_1 0$, it follows that

$$\log \left(1 + \frac{x^2}{\omega^2 \tan^2(k\pi/\omega)} \right) = {}_1 \log \left(1 + \frac{x^2}{k^2\pi^2} \right), \quad x \in R,$$

and so $\eta_n = {}_1 0$ for all finite $n \in N$. Then it follows from Theorem 6.5 that there exists an infinitely large natural number $v \leq [(\omega-1)/2]$ such that $\eta_n = {}_1 0$ for all $n \leq v$.

Observing that

$$\log \left(1 + \frac{x^2}{\omega^2 \tan^2(k\pi/\omega)} \right) \geq 0 \quad \text{for all } 1 \leq k \leq [(\omega-1)/2],$$

we obtain that

$$0 \leq \eta_{[(\omega-1)/2]} \leq \eta_v + \sum_{k=v+1}^{[(\omega-1)/2]} \log \left(1 + \frac{x^2}{k^2\pi^2} \right) = {}_1 \sum_{k=v+1}^{[(\omega-1)/2]} \log \left(1 + \frac{x^2}{k^2\pi^2} \right).$$

From Cauchy's criterion in the form of Theorem 6.4 it follows, however, that $\sum_{k=v+1}^{[(\omega-1)/2]} \log(1 + (x^2/k^2\pi^2)) = {}_1 0$ for all $x \in R$, and so we obtain that $\eta_{[(\omega-1)/2]} = {}_1 0$. Finally, the lemma follows from the continuity of the log-function.

In order to complete the proof observe that from (9.4) and (9.7) it follows that for all standard $x \in R$ we have

$$\sinh x = {}_1 x \prod_{k=1}^{[(\omega-1)/2]} \left(1 + \frac{x^2}{\omega^2 \tan^2(k\pi/\omega)} \right), \quad \omega \in {}^*N - N,$$

and so by taking standard parts using Lemma (9.9) we obtain finally that

$$\sinh x = x \prod_{k=1}^{\infty} \left(1 + \frac{x^2}{k^2\pi^2} \right) \text{ for all } x \in R.$$

From the latter formula the product formula can be obtained by using the uniqueness theorem for analytic functions. In this connection it is not without interest to remark that a slight extension of the argument presented above will give

the result for all complex $z \neq \pm k\pi$, $k = 0, 1, 2, \dots$. We shall leave it to the reader to verify this.

10. Nonmeasurable functions. In this final section of the present paper we shall present a simple example of a function which is not measurable in the sense of Lebesgue. The construction or rather the definition of the example will be based on the theory of infinitely large and infinitely small numbers.

In a previous paper [5], the present author already defined such a function. It involved some nontrivial properties of the sine-function in *R . We shall follow here another idea.

Let $\omega \in {}^*N - N$ be an infinitely large natural number. Then by Theorem 3.10 the following function is internal.

$$(10.1) \quad \phi(x) = [2^\omega x] - 2[2^{\omega-1}x], \quad x \in {}^*R.$$

The internal function ϕ is obviously periodic modulo one. It can also be defined as the ω th coefficient of the dyadic expansion of $x - [x]$ ($x \in {}^*R$), and so it takes on only the values 0 and 1.

By f we shall denote the restriction of ϕ to the set of standard real numbers R of *R . Then the following result holds.

THEOREM 10.2. *The real function $f(x) = [2^\omega x] - 2[2^{\omega-1}x]$, $x \in R$, is not measurable in the sense of Lebesgue.*

Proof. Observe that f has the following properties. (i) For every (standard) dyadic number d , $0 \leq d \leq 1$, $f(d) = 0$. (ii) Every dyadic number d , $0 \leq d \leq 1$, is a period of f , that is, $f(x + d) = f(x)$ for all $x \in R$. (iii) For all x , $0 \leq x \leq 1$, we have $f(1 - x) = 1 - f(x)$ provided x is not dyadic. (i) and (ii) follow immediately from the fact that since ω is infinitely large, $2^\omega d$ and $2^{\omega-1}d$ are natural numbers for all standard dyadic numbers d , $0 \leq d \leq 1$. (iii) follows from the fact that $f(x)$ is the ω th coefficient of the dyadic expansion of x in *R . We shall now assume that f is a measurable function. We shall have to use the following well-known result.

(10.3) *A measurable function which has arbitrarily small periods is equal to a constant almost everywhere.*

From the assumption that f is measurable, property (ii) of f , (10.3), and the fact that f takes on only the values 0 and 1 it follows that $f = 0$ a.e. or $f = 1$ a.e.

Let $A = \{x: 0 \leq x \leq 1 \text{ and } f(x) < 1/2\}$. Then A is a measurable set, and the characteristic function of A has all the dyadic number as periods, and so, by (10.3), $m(A) = 0$ or $m(A) = 1$, where m denotes the Lebesgue measure. Consider now also the set $B = \{x: 0 \leq x \leq 1 \text{ and } f(x) > \frac{1}{2}\}$. Then properties (ii) and (iii) of f imply that if x is not dyadic $0 < x < 1$, then $x \in A$ if and only if $1 - x \in B$. Thus

the set A_0 of non-dyadic points of A and the set B_0 of nondyadic points of B are symmetric with respect to the point $1/2$. Since the set of dyadic points is countable its Lebesgue measure is zero and so $m(A) = m(A_0) = m(B_0) = m(B)$. Then $A_0 \cap B_0 = \emptyset$, $m(A_0 \cup B_0) \leq 1$, $m(A_0) = m(B_0)$ and $m(A_0) = 0$ or $m(A_0) = 1$ imply that $m(A_0) = m(B_0) = 0$. Hence, $f(x) = 1/2$ a.e., which contradicts the fact that f does not take on the value $1/2$. We conclude that f is not measurable in the sense of Lebesgue and the proof is finished.

Work on this paper was also supported in part by Grant No. GP-7691 from the National Science Foundation.

References

1. A. R. Bernstein and A. Robinson, Solution of an invariant subspace problem of K. T. Smith and P. R. Halmos, *Pacific J. Math.*, 16 (1966) 421–431.
2. A. R. Bernstein, Invariant subspaces of polynomially compact operators on a Banach space, *Pacific J. Math.*, 21 (1967) 445–464.
3. W. A. J. Luxemburg, *Non-Standard Analysis*. Lectures on A. Robinson's theory of infinitesimals and infinitely large numbers, Pasadena (1962) and revised edition (1964).
4. ———, Two applications of the method of construction by ultrapowers, *Bull. Amer. Math. Soc.*, (2) 68 (1962) 416–419.
5. ———, Addendum to, "On the measurability of a function which occurs in a paper by A. C. Zaanen", *Proc. Roy. Acad. Sci., Amsterdam*, A66 (1963) 587–590.
6. ———, *Applications of Model Theory to Algebra, Analysis and Probability Theory*. Proceedings of an International Symposium on Nonstandard Analysis, Holt-Rinehart and Winston, New York, 1969.
7. A. Robinson, Non-standard analysis, *Proc. Roy. Acad. Sci., Amsterdam*, A64 (1961) 432–440.
8. ———, *Non-standard Analysis*, *Studies in Logic and the Foundations of Mathematics*, North-Holland, Amsterdam, 1966.
9. ———, On generalized limits and linear functionals, *Pacific J. Math.*, 14 (1964) 269–283.
10. ———, On the theory of normal families, *Acta Philos. Fenn.*, 18 (R. Nevanlinna anniversary volume) (1965) 159–184.
11. ———, On some applications of model theory to algebra and analysis, *Rend. Mat. e Appl.*, 25 (1966) 1–31.
12. ———, A new approach to the theory of algebraic numbers, *Atti Accad. Naz. Lincei Rend.*, (8) 40 (1966) 222–225, 770–774.
13. ———, Nonstandard theory of Dedekind rings, *Proc. Roy. Acad. Sci., Amsterdam*, A70 (1967) 444–452.
14. ———, Nonstandard arithmetic, *Bull. Amer. Math. Soc.*, 73 (1967) 818–843.
15. R. T. Taylor, *Invariant Subspaces in Hilbert and Normed Spaces*, Ph. D. Thesis, Calif. Institute of Technology, Pasadena, 1968.

RECURSIVE FUNCTIONS AND HIERARCHIES

HILARY PUTNAM, Harvard University

1. Introduction. Everyone knows that modern computing machines can solve many mathematical problems. It is true that they do not exercise “ingenuity” in solving these problems: the operator has to give the machine exact rules of procedure, and the machine can only proceed by following these rules—the so-called “program”—step by step. Moreover, the rules of procedure must themselves be framed in terms of a highly limited vocabulary—a vocabulary corresponding, in a certain sense, to the elementary operations that the machine is able to carry out. But these limitations are compensated for, to some extent, by the immense speed of the machines which enables them to duplicate many lifetimes of human computation in a few minutes.

When we attempt to study what such machines can or cannot in principle accomplish, we obviously have to make certain idealizations. It goes without saying that we ignore wear and tear on the machines, mechanical malfunctions, etc.—i.e., that what we consider are abstract machines. An important part of the idealization is this: we pretend that the machines have available a *potentially infinite external memory*.

To suppose that the external memory is potentially infinite is to suppose that there is no limit to the number of filing cabinets full of IBM cards (or whatever) that may be placed next to the machine for the machine to use for data storage—and also that the machine has the right to “demand” more such external memory space if it is necessary for it to store more data. In short, the machine is supposed to be able to store and recover arbitrary finite amounts of data.

When we make these idealizations—no wear and tear on the machine (i.e., the machine is “immortal”), no mechanical malfunctions, potentially infinite external memory—something really remarkable happens. It turns out that the problem solving capacity of any one of the standard computing machines is identical with the problem solving capacity of any of the others. Of course, one machine may solve a problem quickly while another solves it more slowly, depending on the design of the machine, but if one machine can solve a problem *P* at all, *any* standard computing machine can solve it (given enough time).

Moreover, it turns out that certain very simple machines—very simple to describe mathematically—are “universal” in the sense that they can solve all of the problems that any computing machine can solve. These machines are ideal for the purpose of studying mathematically just what is and what is not computable in principle.

Instead of beginning our discussion of computability with a mention of computing machines, we might as well have started with the notion of an algorithm. Indeed, a computing machine is merely a device for carrying out algorithms (a “universal”

machine is one that can be programmed to carry out an arbitrary algorithm). The notion of algorithmic computability has been formalized in several different ways: the **basic result** on which the subject of recursive function theory is founded is that the class of functions computable in any of these different formalizations is exactly the same. (The use of the term “basic result” to cover the coextensiveness of the different formal definitions of algorithmic computability is due to Rogers.

The thesis that any function that can be computed by following out the steps of an algorithm belongs to the class of functions that are computable by using Gödel-Kleene systems of equations, or Church λ -calculi, or Turing machines (in view of the **basic result** it does not matter which of these notions of algorithmic computability we employ to define the class in question), is not a precise mathematical statement because it relates the essentially imprecise term “algorithm” to a precisely defined term (e.g., “function computable by a Turing machine”). This thesis—**Church’s Thesis**, as it has come to be called—nevertheless plays an important role in motivating the study of the class of functions in question and of the formalisms for computability just mentioned. The **basic result** constitutes part of the “evidence” for Church’s Thesis: the fact with which we began this lecture, the fact that the functions computable (or the “problems solvable”) using any one of the standard computing machines are exactly the functions in this class, constitutes a further piece of evidence for Church’s thesis.

We shall not review the definition of a Turing machine here. We shall, however, use the following facts about Turing machines in this lecture: (1) Turing machines accept as input finite strings in a fixed finite alphabet. They store data and carry out computations by scanning the input and by printing finite strings in the same alphabet, if necessary erasing the original input in the course of the computation. (In fact, they do all this subject to the restriction that they may only scan one symbol at a time, and they may only move right or left one space at a time.) (2) When a Turing machine has finished its computation it “halts” (i.e., it comes to rest in a distinguished state, the “rest state”). The answer to the original problem (assuming the machine gave an answer) is then to be found on the tape on which the machine prints symbols. For example, if the answer is a number, then that number (in, say, unary or binary notation) will be printed on the tape, flanked, if necessary, by identifying marks. (There may be other “junk” left on the tape, but the answer will always be distinguished by some such device as the device of an identifying mark at the beginning and end.) (3) We assume that the machine has an infinite two-way tape on which to compute: this is the special form that the assumption of a “potentially infinite external memory” takes in the case of this particular kind of computing device. However, the original input—what is printed on the tape at the start of a computation—is always required to be finite. (4) The class of functions computable by Turing machines with a fixed n -sign alphabet is independent of the size n of the alphabet, provided $n \geq 2$. Here “function” means “function from natural numbers

to natural numbers'', and ''computable'' means that if the machine is given the arguments of the function as input (i.e., these are printed on its tape in unary notation), then it will always come to rest in a finite time (which may not be ''easily calculable'' in advance given the arguments), and the correct value of the function for the arguments given as input will be found printed on its tape. (Also, it is required that this value be the *only* number printed on its tape when it halts, or the only number flanked by the distinguishing mark, if a distinguishing mark is used.) (5) One can describe all possible Turing machines over a fixed n -sign alphabet by means of certain canonical expressions. These expressions can themselves be coded in the n -sign alphabet (they can even be coded as natural numbers), and a suitable Turing machine with that alphabet can even enumerate all of the canonical descriptions, in that coding. Henceforth, we shall imagine that a fixed n -sign alphabet (n equal to or larger than 2) has been picked, and that a fixed coding of the canonical descriptions of the Turing machines over that alphabet has been picked, and that the coded canonical descriptions, in the order in which a certain Turing machine lists them, are C_1, C_2, C_3, \dots . We shall also use the expression T_i for the Turing machine whose description is C_i ($i = 1, 2, 3, \dots$): thus T_1, T_2, T_3, \dots is a list of *all* Turing machines (over the fixed alphabet).

Let A be an infinite set of ''yes-no'' questions, or an infinite set of numerical questions (questions to which the answer is a natural number). (We assume the questions in the set A to be expressions in the fixed n -sign alphabet.) If a Turing machine can be programmed to write down the correct answer (''yes'' or ''no'', or the appropriate natural number) whenever the machine is given a question belonging to the set A as input, we say the **decision problem** for A is solvable. Otherwise, we say that this problem is unsolvable.

Historically, the first interest of recursive function theorists was in showing that various decision problems were unsolvable. ''Negative solutions''—proofs of unsolvability—were the hallmark of the field. While recursive function theory today covers many other topics, and while its methods are so widespread and so integrated with other branches of mathematical logic that it is difficult to say where recursive function theory ends and, say, set theory begins, the study of unsolvable problems remains an important special topic.

Proofs that a problem is unsolvable by specified means are not new in the history of mathematics. Omar Khayam had already conjectured that certain algebraic equations are not solvable by means of radicals (indeed, he thought mistakenly that this already happens at degree three), and Galois showed that in fact the general fifth degree equation is not solvable. Showing that a problem is not solvable by means of Turing machines is importantly different from showing that it is not solvable by means of radicals, however. If an equation is not solvable by means of radicals, we can still express the solution in other ways, and there are perfectly good methods for obtaining this solution to as many decimal places as we wish. But if a decision problem is not solvable by means of Turing machines, we are stuck—there are no

more general uniform procedures. In this sense, the unsolvability proofs of recursive function theory are the first proofs of **absolute unsolvability** in human history. (Of course, the statement that there are no more general uniform procedures than Turing machines is a form of Church's thesis; thus this last remark presupposes that thesis.)

2. The basic concepts of recursive function theory. The functions computable by Turing machines are called **recursive**. f is recursive if and only if the decision problem for questions of the form " $f(n) = ?$ " is solvable.

A set of natural numbers is called **recursive** if and only if the characteristic function of the set is recursive. A is recursive if and only if the decision problem for questions of the form "Does n belong to A ?" is solvable.

An important distinction is the distinction between recursiveness and **recursive enumerability**. A set A of natural numbers is called **recursively enumerable** if A is either empty or the set of values taken on by some recursive function. In terms of Turing machines what this means is just that it is possible to program a Turing machine so that the machine will print out all and only the natural numbers in the set A . (We count the machine as having "printed out" a natural number just in case that natural number appears on its tape with the fixed distinguishing mark at the beginning and end at some stage. The requirement that the distinguishing mark appear at the beginning and end enables us to distinguish what the machine actually "prints out" from the "data" that the machine stores and from its "calculations".) Of course, if A is an infinite recursively enumerable set, we must allow the machine to run forever so that it can print out all the members of the set A . In short, the machine **generates** the members of A as an infinite list (not necessarily in any natural order).

If the "basic result" of recursive function theory is the coextensiveness of the different definitions of algorithmic computability, then the following result might be called the *Hauptsatz* of the subject: *There exists a recursively enumerable set of natural numbers which is not recursive.*

Let us give an example. Imagine a machine M which writes down all finite sequences of well-formed formulas of quantification theory (i.e., of elementary formal logic, including symbols for relations and symbols for the quantifiers "for all x " and "there exists an x such that"). Such a Turing machine obviously exists. If now we complicate the machine by having it test each finite sequence, to check if that finite sequence is or is not a *proof* on the basis of the rules of some particular standard system of quantification theory, and to print the distinguishing mark before and after the last line (i.e., the last formula in the finite sequence) if and only if the finite sequence is a proof, then what will the result be? Clearly the machine M will print out *only* theorems of quantification theory; and since every theorem is (by definition) the last line of some proof, M will print out each theorem of quantification theory. Thus the set TH of all theorems of quantification theory is a recursively enumerable set.

However, Church proved in 1936 that the decision problem for TH is unsolvable. (By the “decision problem for TH ” we mean, of course, the decision problem for questions of the form “does F belong to TH ”, where F is a well-formed formula of quantification theory.) Thus TH is not a recursive set.

(By the above argument, the set of theorems of any formal system is a recursively enumerable set; but it is not a recursive set unless the formal system has a solvable decision problem.)

This example has the defect that TH is a set of formulas and not a set of natural numbers, but this is inessential: it is trivial to code formulas of any formal system as natural numbers, and thus one can obtain a set of integers with the same property as TH : the property of being recursively enumerable but not recursive. (It is only necessary that the code employed be “mechanical”, that is, that the transition between a formula and its number, or between a number and the formula it encodes, be capable of being carried out by a Turing machine. But any code that would naturally occur to one has this property.)

In addition to the concepts just defined, the concept of a **partial recursive function** plays an important role in recursive function theory. A partial function is a function whose domain is a subset of the natural numbers and whose range is a subset of the natural numbers. (Thus ordinary “total” functions are also partial functions, and the “totally undefined function”—i.e., the “empty” function—is a partial function.) If the domain of a partial function f does not include a natural number n , then f is said to be **undefined** on n .

A partial function f is called **partial recursive** just in case a Turing machine can be programmed to print out all and only the true statements of the form “ $f(n) = m$ ”. This is equivalent to requiring that it be possible to program a Turing machine so that (1) it answers correctly all questions of the form “ $f(n) = ?$ ”, where n is a natural number on which f is defined; and (2) the machine runs forever (without printing out an answer) when confronted with the question “ $f(n) = ?$ ” if n is a natural number on which f is not defined.

The domain of a partial recursive function need not be recursive, but it is clear from the above definition that it must always be a recursively enumerable set.

3. Unsolvable problems. As already remarked, the first interest of recursive function theorists was in showing that various decision problems were unsolvable. Once one has succeeded in showing that one decision problem is unsolvable—say, the decision problem for questions of the form “does n belong to A ”, where A is some fixed set of natural numbers—it is not difficult, in general, to go on and show that many other decision problems are unsolvable. For example, suppose that I can show that the decision problem for the set A would be solvable if the decision problem for a certain set B were solvable. I might do this, for example, by showing that there is a recursive function f such that, for all n , $n \in A$ if and only if $f(n) \in B$. Since I already know that the decision problem for A is unsolvable, it follows that the

decision problem for B is likewise unsolvable. In this way a great many mathematical problems have been shown to be unsolvable. The decision problem mentioned in the preceding section—the decision problem for the set TH was shown to be unsolvable by this method.

But this method presupposes that some problem other than the one whose unsolvability one wishes to establish is already known to be unsolvable. How was the *first* problem shown to be unsolvable?

The answer turns on a basic theorem of recursive function theory—the enumeration theorem. This will now be sketched.

For $i = 1, 2, 3, \dots$, let W_i be the set of all natural numbers T_i ever prints out. (Recall that T_i is the i th Turing machine in the enumeration described in §1.) Since every recursively enumerable set is the set of all integers printed out by some Turing machine, it follows that W_1, W_2, W_3, \dots is an enumeration of all recursively enumerable sets. Let $J(x, y)$ be any one-one recursive function mapping pairs of integers onto integers (for the sake of definiteness, take $J(x, y) = 2^x 3^y$). Then the **enumeration theorem** states that the set of integers $J(x, y)$ such that $x \in W_y$ is recursively enumerable.

Whenever $\{J(x, y) \mid P(x, y)\}$ is recursively enumerable (where $P(x, y)$ is a two-place predicate of natural numbers), we say that the predicate P in question is a recursively enumerable two-place predicate. In this terminology, what the enumeration theorem says is that the two-place predicate " $x \in W_y$ " is a recursively enumerable predicate.

The proof is a direct construction of a Turing machine which lists all possible computations of all possible Turing machines and examines these computations to discover all truths of the form " $x \in W_y$ ".

A second theorem of recursive function theory is this: a set A is recursive if and only if A and \bar{A} are both recursively enumerable. It is clear that if A is recursive (i.e., a Turing machine can answer all questions of the form "does n belong to A "), then A and \bar{A} are both recursively enumerable. To verify the converse, let M_1 be a machine which generates the members of A and let M_2 be a machine which generates the members of \bar{A} . Then it is possible to construct a Turing machine M_3 which operates as follows: M_3 lists all possible computations of both M_1 and M_2 (in a suitable encoding). Suppose someone gives M_3 a question of the form "does n belong to A ". Then M_3 lists all computations of M_1 and M_2 , as we just said, and scans the "print out" of M_1 and the "print out" of M_2 until one or the other machine prints out the given integer n . If M_1 ever prints out n , then M_3 prints out "yes" and halts as soon as it comes to that computation of M_1 ; if M_2 ever prints out n , then M_3 prints out "no" and halts as soon as it comes to that computation. Since n must belong to either A or \bar{A} (and not to both), either M_1 or M_2 must eventually print out n , and so M_3 must eventually answer the question.

Let $D = \{x \mid x \in W_x\}$. Since $x \in D$ if and only if $J(x, x)$ belongs to the set $\{J(x, y) \mid x \in W_y\}$, and this latter set is recursively enumerable, it is easily seen that D

is recursively enumerable (just program a Turing machine to scan all computations of the machine that enumerates $\{J(x, y) \mid x \in W_y\}$ and to test the print out of that machine to see if the number printed out is of the form $J(x, x)$ —this is easy to determine effectively. Instruct the machine that it is to print out those numbers that it examines that pass this test—i.e., those numbers that are (1) printed out by the original machine, and (2) of the form $J(x, x)$.) *But D is not recursive.* For suppose it were. Then \bar{D} would be recursively enumerable—i.e., there would be an integer k such that $\bar{D} = W_k$. Then,

$k \in \bar{D}$ iff k does not belong to D
 iff it is not the case that $k \in W_k$
 iff it is not the case that $k \in \bar{D}$
 —a contradiction.

The set D —i.e., the set of all x such that $x \in W_x$ —is thus an example (in fact, the basic example) of a set which is recursively enumerable but not recursive, and the decision problem for D is the basic example of an unsolvable problem. The unsolvability of the decision problem for D was the basic tool used in establishing the unsolvability of the decision problem for quantification theory, the decision problem for number theory, and many other problems.

4. Relative unsolvability. Let A and B be nonrecursive sets of natural numbers. In other words, let the decision problem for A and the decision problem for B (i.e., the decision problem for questions of the form “does $n \in A$ ” and the decision problem for questions of the form “does $n \in B$ ”) both be unsolvable problems. It may still happen that one of these two decision problems is “more unsolvable” than the other; and the study of **relative** unsolvability—of “more” and “less” unsolvability—has become a big topic in recursive function theory.

What is meant by saying that the decision problem for B is more unsolvable than the decision problem for A is that

- (1) If B were recursive, then A would be recursive; but
- (2) It is not the case that B would be recursive if A were.

But how can we give (1) and (2) a precise meaning?

What we want to formalize is the following notion: that a Turing machine could answer all questions of the form “does $n \in A$ ” correctly, provided the machine had access to an “oracle” for B —that is, to a device which answered correctly all questions of the form “does $n \in B$ ”. One simple way of doing this is the following: modify the notion of a Turing machine so that a Turing machine can scan and move two tapes instead of one. Let one of these tapes contain at the beginning of the computation only the particular question “does $n \in A$ ” which the machine is supposed to answer. This is the tape on which the machine will perform its calculations and

print its final answer. Let the other tape contain all true statements of the form " $n \in B$ " and "it is not the case that $n \in B$ ". (Apart from the contents of the second tape, the machine is still describable by a finite canonical description, just like an ordinary Turing machine. Moreover, the machine can scan either tape only one square at a time, and can move either tape only one square at a time. Thus it can, in fact, only use *finitely* much information about the set B in any one computation, although it has all of the set B potentially available to it). Call such a machine a B -machine. Then call the set A **recursive relative to B** (or "recursive in B ") just in case it is possible to program a B -machine to answer correctly an arbitrary question of the form "does $n \in A$ ". Finally, give (1) and (2) a precise meaning by saying:

(3) A is **more unsolvable than B** just in case B is recursive in A and A is not recursive in B .

Another important case is the case in which A and B are recursive in each other. In this case we say that A and B are **Turing equivalent**.

We now introduce the notion of a **degree of unsolvability**:

(4) A degree of unsolvability is the class of all sets of natural numbers which are Turing equivalent to some given set A .

Here we are exploiting the fact that Turing equivalence is an equivalence relation; the degrees of unsolvability are just the equivalence classes of this relation.

(5) If d_1 is the degree of unsolvability of a set A (i.e., A is an element of the equivalence class d_1), and d_2 is the degree of unsolvability of a set B , then we say that $d_1 \leq d_2$ just in case A is recursive in B . (It is easily checked that this is well defined, i.e., that whether $d_1 \leq d_2$ or not is independent of the particular choice of the representatives A, B .)

(6) The degree of the recursive sets (this is the same as the class of all recursive sets) is denoted by the symbol 0 . Since a recursive set is (trivially) recursive relative to every set, $0 \leq d$ for all degrees d .

The structure of the system of degrees of unsolvability is extremely complicated. Post and Kleene long ago showed that there are **incomparable** degrees: i.e., the ordering of degrees is not linear. A famous result of recursive function theory is this: *there are incomparable r.e. degrees*, i.e., degrees d_1, d_2 of *recursively enumerable* ("r.e.") sets such that neither $d_1 \leq d_2$ nor $d_2 \leq d_1$ holds. This was proved in 1957 by Richard Friedberg in the course of solving "Post's Problem" (whether all non-recursive r.e. sets have the same degree).

Let d_1 and d_2 be degrees of sets A and B . Let

$$C = \{J(x, y) \mid x \in A \text{ \& } y \in B\}.$$

(C is a kind of "recursive cartesian product" of A and B .) It is easily seen that the degree of C is a least upper bound on the degrees d_1, d_2 ; however, it has been proved that greatest lower bounds do not in general exist. Thus the partial ordering of degrees is not a lattice, but only an upper semilattice. A great deal of work has gone

into discovering the structure of this upper semilattice, and of the subsystem of r.e. degrees. The structure is extremely messy. It has recently been shown, in fact, that an arbitrary countable partial ordering can be imbedded in the ordering of degrees of unsolvability

5. Relative recursive enumerability and the jump operator. Let A, B be sets of natural numbers. If a B -machine can be programmed to print out exactly the set A (i.e., the membership of A is printed out in some order, not necessarily a natural order, on the machine's calculation tape, as the machine goes on running forever) then the set A is said to be **recursively enumerable relative to B** (or "r.e. in B ").

Intuitively, if A is r.e. in B , then we may say that "if B were **recursive**, then A would be **recursively enumerable**"; or "a Turing machine could generate the set A , provided it had access to an oracle for the set B ." The relation "r.e. in", unlike the relation "recursive in" is not transitive. A set A is r.e. in a recursive set just in case A is itself r.e. (as is easily seen); but a set which is r.e. in an r.e. set need not be r.e. To show this let $K = \{J(x, y) \mid x \in W_y\}$. Then, for any k , $W_k = \{x \mid J(x, k) \in K\}$, and hence if one had access to an "oracle" for K , one could find out if an arbitrary integer n belongs to W_k by asking the oracle the single question "does $2^n 3^k$ belong to K ". So it is easy to program a K machine to answer correctly all questions of the form "does $n \in W_k$ ", for any given k . Therefore, K is an r.e. set in which every other r.e. set is recursive; such an r.e. set is called a **complete** r.e. set.

(If there exists a *recursive* function f such that, for arbitrary n , $n \in A$ if and only if $f(n) \in B$, then we say that A is "many-one reducible" to B . Many-one reducibility is only one of a number of different types of reducibilities stronger than Turing reducibility that have been studied. The above proof that every r.e. set is recursive in K obviously gives the stronger result that every r.e. set is many-one reducible to K .)

We saw in §3 that $D = \{x \mid x \in W_x\} = \{x \mid J(x, x) \in K\}$ is not recursive. But D is obviously recursive in K ; so K also is not recursive.

Now, let C_1^B, C_2^B, \dots be any natural enumeration of canonical descriptions of all possible B -machines. (The descriptions do not include the contents of the B -tape; thus the descriptions are finite, and, in fact, very similar to the descriptions C_1, C_2, \dots of the ordinary Turing machines.) Let T_1^B, T_2^B, \dots be the B -machines whose descriptions are (respectively) C_1^B, C_2^B, \dots . Let W_i^B , for $i = 1, 2, 3, \dots$ be the set of all natural numbers ever printed out by T_i^B . Then the predicate $x \in W_y^B$ can be shown to be recursively enumerable in B (i.e., $\{J(x, y) \mid x \in W_y^B\}$ is r.e. in B).—This is called the **relativized enumeration theorem**: the proof is a direct construction of a B -machine which writes down all computations of all possible B -machines and examines them in order to list all truths of the form " $x \in W_y^B$ ". Let.

$$K^B = \{J(x, y) \mid x \in W_y^B\}.$$

By the relativized enumeration theorem, K^B is r.e. in B . Moreover, every set which is r.e. in B is many-one reducible to K^B and hence recursive in K^B (by just the argument

that we gave before to show that all r.e. sets are recursive in K). K^B is a **complete B -r.e. set**. But $D^B = \{x \mid x \in W_x^B\}$ is not recursive in B , and so neither is K^B . Let us now take $B = K$, i.e., let us consider K^K . Since all sets which are recursively enumerable are recursive in K , and K^K is not recursive in K , we see that K^K is not r.e. However, K^K is r.e. in K ; in fact, it is the complete K -r.e. set, as we just saw. Thus we have proved that a set which is r.e. in an r.e. set need not be r.e.

Generalizing the argument, we see that if B is any set, then K^B is a complete B -r.e. set, and

K^{K^B} is a set which is r.e. in K^B but not
recursive in K^B , and hence not B -r.e.

Thus

K^{K^B} is an example of a set which is r.e.
in a set which is r.e. in B , but not itself r.e. in B .

If d is the degree of a set B , we define d' (the "jump" of d) to be the degree of a complete B -r.e. set. (If B is a set of natural numbers, then the notation B' is also used for the set K^B . In this notation K^K is denoted as K' , $K^{K'}$ as K'' , etc.) Since any two complete B -r.e. sets are recursive in each other, it is clear that d' is well defined (i.e., d' is independent of the choice of the particular representative B from the equivalence-class d and of the particular complete B -r.e. set). Since 0 is the degree of the recursive sets, $0'$ is the degree of the complete r.e. set; if d is an r.e. degree, then $0 \leq d \leq 0'$. Post's problem (referred to before) asked if there were any r.e. degrees other than 0 and $0'$; Friedberg succeeded in showing that there are infinitely many, and even that there exists an infinite family of pairwise incomparable r.e. degrees).

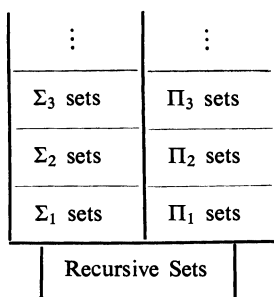
6. The arithmetical hierarchy and the Kleene-Post representation theorem. Suppose we wished to classify the sets of integers most commonly encountered in mathematics. One way which naturally comes to mind is to classify them according to the structure of their definitions. Thus a set which can be defined by a definition of the form $\{n \mid (x)Rx_n\}$, would be in one "box" of the classification; a set which can be defined by a definition of the form $\{n \mid (Ex)(y)Rxy_n\}$, where R is a recursive 3-place predicate, would be in a different box, and so on. The whole structure looks like this:

\vdots	\vdots
$(x_1)(Ex_2)(x_3)Rx_1x_2x_3n$	$(Ex_1)(x_2)(Ex_3)Rx_1x_2x_3n$
$(x_1)(Ex_2)Rx_1x_2n$	$(Ex_1)(x_2)Rx_1x_2n$
$(x_1)Rx_1n$	$(Ex_1)Rx_1n$
<div style="border: 1px solid black; padding: 2px; display: inline-block;">R</div>	

THE ARITHMETICAL HIERARCHY — FORM OF DEFINING PREDICATE

(R is here used as a special variable for recursive predicates)

The symbols “ (x) ” and “ (Ex) ” are just the universal quantifier and the existential quantifier of formal logic. (“ (x) ” is read “for every x ”, and “ (Ex) ” is read “there is an x such that”). Thus $\{n \mid (x_1) (Ex_2) Rx_1x_2n\}$, for example, is simply the set of all natural numbers n which are such that no matter what natural number x_1 one selects, it is possible to find a natural number x_2 , possibly depending on both n and on x_1 , such that the three numbers x_1, x_2, n stand in the three-term recursive relation Rx_1, x_2, n .) A predicate which consists of n alternating quantifiers of which the first is universal and then a recursive predicate is called a Π_n predicate; a predicate which consists of n alternating quantifiers of which the first is existential and then a recursive predicate is called a Σ_n predicate. Also, a set which can be defined by a Σ_n (respectively Π_n) predicate is called a Σ_n (respectively Π_n) set. The “ Σ ” “ Π ” notation is obviously motivated by the traditional analogy between the existential and universal quantifiers and the operations of summation and product respectively. Using this notation, we can redraw the diagram thus:



THE ARITHMETICAL HIERARCHY

In view of the elementary equivalences

$$\neg(x)R \equiv (Ex)\neg R$$

$$\neg(Ex)R \equiv (x)\neg R$$

and the fact that the complement of a recursive predicate is recursive, we see that on each level, the sets in one box are the complements of the sets in the box immediately to the side. Also a recursive set, say $\{n \mid Rn\}$, can be defined by a Σ_n predicate, for arbitrary n , e.g., as $\{n \mid (Ex_1) (x_2) (Ex_3) \cdots (Qx_n) (Rn \& x_1 = x_1 \& x_2 = x_2 \& \cdots \& x_n = x_n)\}$, where “ (Qx_n) ” is an existential (universal) quantifier if n is odd (even), and by a similar trick, by a Π_n predicate; and by the same argument, every set that can be defined by a Σ_n or Π_n predicate can be defined by a Σ_{n+k} and by a Π_{n+k} predicate for all $k \geq 1$. Thus the levels are cumulative: every set of a given box belongs to both boxes on any higher level.

The *Hierarchy Theorem* of Kleene asserts that the classification is indeed a

hierarchy: that is (1) Both boxes in every level above the base contain sets not found at lower levels; and (2) Every box contains a set not found in the box immediately to the side.

For example, the r.e. sets are exactly the Σ_1 sets. The set K therefore lies in the box of Σ_1 sets; but it does not lie in the box of Π_1 sets, because the Π_1 sets are just the sets with r.e. complements, and K does not have an r.e. complement. Similarly, K^K lies in the box of Σ_2 sets, but not in the box of Σ_1 sets nor in the box of Π_1 sets.

Why did we leave out of our classification sets whose defining predicates have the form $(x_1)(x_2)Rx_1x_2n$?

The answer is that

$$\{n \mid (x_1)(x_2)Rx_1x_2n\} = \{n \mid (x)R(K(x), L(x), n)\},$$

where, for natural numbers x , we define

$K(x)$ = the exponent of 2 in the prime factorization of x ,

$L(x)$ = the exponent of 3 in the prime factorization of x .

Since the predicate $R(K(x), L(x), n)$ is recursive whenever R is, this shows that any set that can be defined with two successive universal quantifiers followed by a recursive predicate can be defined with just *one* universal quantifier. In the same way, any number of *successive* quantifiers of like quality—all existential or all universal—can be “contracted”. Thus, in applying successive quantifiers to a recursive predicate, we can assume without loss of generality that the quantifiers alternate in quality.

According to a theorem of formal logic, any predicate that can be defined starting from predicates P_1, P_2, \dots, P_k by using the familiar truth-functions (“or”, “and”, “not”, etc.) and quantifiers of formal logic can be defined by a “normal form” predicate—one consisting of a “prefix” of a string of quantifiers followed by a “matrix” which is a predicate constructed out of the given predicates using truth-functions alone. For example:

$$\begin{aligned} - ((x_1)Px_1nx_1 \vee (Ex_2)Qx_2n) &\equiv - (x_1)Px_1nx_1 \& - (Ex_2)Qx_2n \\ &\equiv (Ex_1) - Px_1nx_1 \& (x_2) - Qx_2n \\ &\equiv (Ex_1)(- Px_1nx_1 \& (x_2) - Qx_2n) \\ &\equiv (Ex_1)(x_2)(- Px_1nx_1 \& - Qx_2n). \end{aligned}$$

But a predicate which is constructed out of recursive predicates using truth-functions alone is recursive (computability is obviously closed under truth-functions); thus any set that can be defined starting from recursive predicates by using truth-functions and quantifiers can be defined by a predicate consisting of quantifiers followed by a *recursive* predicate. If we then “contract” any successive quantifiers that happen to be of the same quality (all existential or all universal), in the fashion

explained above, we get a predicate of one of the forms Σ_n or Π_n (unless there are no quantifiers at all, in which case we have a recursive predicate). Thus every such set belongs to our hierarchy. In particular, every set that can be defined in **first order arithmetic** belongs to our hierarchy. ("First order arithmetic" is a formal system whose variables range over natural numbers and whose undefined predicates are just $x = y + z$ and $x = y \cdot z$.) Since every recursive predicate can be defined from the predicates $x = y + z$ and $x = y \cdot z$ using truth-functions and quantifiers (the basic trick is a way of coding finite sequences of natural numbers which was contributed by Gödel in his epochal paper on undecidable sentences), the converse is also true: the Arithmetical Hierarchy is a classification of *exactly* the predicates that can be defined in first order arithmetic.

The second way of classifying sets of natural numbers that naturally comes to mind, at least if one is a recursive function theorist, is the following:

⋮	⋮
sets r.e. in sets r.e. in an r.e. set	complements of sets r.e. in sets r.e. in an r.e. set
sets r.e. in an r.e. set	complements of sets r.e. in an r.e. set
r.e. sets	complements of r.e. sets
Recursive Sets	

This is an "intrinsic" hierarchy of sets of natural numbers, i.e., one which is not based on the forms of the *definitions* of those sets. Yet a beautiful theorem due to Kleene and Post, the *Kleene-Post Representation Theorem*, asserts that this is exactly the same hierarchy as the Arithmetical Hierarchy. This theorem has a great many interesting consequences. Consider the sequence of sets—

$$K, K^K, K^{K^K}, \dots \quad (\text{or } K, K', K'', \dots)$$

for example. K is a complete r.e. set; K^K is a complete K r.e. set, and hence "complete" for the class of sets r.e. in an r.e. set (i.e., every set which is r.e. in any r.e. set is r.e. in K , and hence recursive in K^K); and, in general, the n th member of the sequence is "complete" for the class of sets in the left hand box of the n th level of the hierarchy depicted (counting the recursive sets as the 0-th level). It follows from the Kleene-Post Representation Theorem that K is also a complete Σ_1 set, K^K a complete Σ_2 set, and, in general, the n th member of the above sequence is a complete Σ_n set (i.e., a Σ_n set in which every Σ_n set is recursive).

We already remarked that even the degrees of r.e. sets do not form a linear ordering, and *a fortiori*, the degrees of arithmetical sets (sets in the above hierarchy) do not form a linear ordering. But the following very simple linear ordering of degrees:

$$0, 0', 0'', 0''', \dots$$

(here 0 again denotes the degree of the recursive sets and ' is the jump operator) at least has the property of dominating all arithmetical degrees—i.e., every arithmetical set has a degree which is \leq all sufficiently large degrees in this sequence.

7. The analytic sets. The predicates $J(x, y) \in W_1, J(x, y) \in W_2, J(x, y) \in W_3, \dots$ are all the recursively enumerable two place predicates. For a two place predicate R is r.e. just in case $\{J(x, y) \mid R_{xy}\}$ is r.e. (by definition: this definition captures the intuitive idea that a two-place predicate is r.e. just in case the ordered pairs $\langle x, y \rangle$ such that R_{xy} can be generated by a Turing machine); hence just in case $\{J(x, y) \mid R_{xy}\}$ is one of the sets W_1, W_2, W_3, \dots in the canonical enumeration of all r.e. sets. Let W_i^2 be the predicate $J(x, y) \in W_i$, for $i = 1, 2, 3, \dots$. Then $W_1^2, W_2^2, W_3^2, \dots$ can be regarded as a canonical enumeration of all two-place r.e. predicates.

It may happen that the predicate W_i^2 is a simple ordering; this can be expressed by the arithmetical condition on the index i :

$$\begin{aligned} (x) (J(x, x) \in W_i) \ \& \\ (x) (y) ((J(x, y) \in W_i \ \& \ J(y, x) \in W_i) \Rightarrow x = y) \ \& \\ (x) (y) (z) ((J(x, y) \in W_i \ \& \ J(y, z) \in W_i) \Rightarrow J(x, z) \in W_i). \end{aligned}$$

—Note that by the theorems of the preceding section the clauses of the form $J(n, m) \in W_i$ can be expressed in the form of an existential quantifier followed by a recursive predicate (because the r.e. set W_i is Σ_1); thus this condition, which we shall abbreviate “SIMPLE (i)” is indeed an arithmetical condition on i .

A very important set, first studied by Markwald, is the set of indices i such that W_i^2 is a well ordering. (This set, the set of indices of r.e. well orderings, was called “ W ” by Markwald; however, it has become standard recently to use “ W ” as the notation for the set of indices of *recursive* well orderings instead). This set cannot be defined by an arithmetical condition, as Markwald proved, but it can be defined by a condition using a quantifier over functions of natural numbers:

i is the index of an r.e. well ordering \equiv

$$\text{SIMPLE}(i) \ \& \ (f) ((x) (J(f(x+1), f(x)) \in W_i) \Rightarrow (Ey) (f(y+1) = f(y)))$$

(i.e., W_i^2 is a well ordering if and only if W_i^2 is a simple ordering and there are no infinite descending chains in the ordering W_i^2).

If we adjoin to the above condition the further clause:

$$(Ex) (y) (z) (J(y, z) \in W_x \equiv J(y, z) \notin W_i),$$

then the definition becomes a definition of the class of recursive well orderings (or, rather, of the corresponding set of indices), for this further clause just says that the predicate W_i^2 has an r.e. complement W_x^2 , and a predicate is recursive just in case it and its complement are both r.e.

The set of indices of r.e. well orderings, and the set of indices of recursive well orderings are examples of **analytic** sets, that is, sets whose definition requires a function quantifier “(f)”, or a string of function quantifiers, followed by a predicate which is built up out of the function variables and ordinary recursive predicates by truth functions, number quantifiers, and composition. The analytic sets have a hierarchy theorem which is directly analogous to the hierarchy theorem for arithmetical sets; but the problem of a representation theorem for these sets is *the* major unsolved problem of hierarchy theory.

We list some elementary facts about function quantifiers and analytical sets for future reference:

(1) Quantification over two-place functions can be reduced to quantification over one-place functions. $(Ef) (x) (Ey) (f(x, y) = 0)$ is equivalent to $(Ef) (x) (Ey) (f(J(x, y)) = 0)$, for example, where the quantifier “(Ef)” is over two-place functions in the formula to the left of “is equivalent to” and over one-place functions in the formula to the right of “is equivalent to”. Similarly, quantification over three and more place functions can be reduced to quantification over one-place functions.

(2) Function quantifiers can be advanced ahead of number quantifiers. Let $\lambda_y f(y)$ be a notation for “the function whose value for arbitrary y is $f(y)$ ”. The advancing of number quantifiers is based on the fact that $(x) (Ef) R(f, x)$ is equivalent to $(Eg) (x) R(\lambda_y g(x, y), x)$. E.g., $(x) (Ef) (f(x + 17) = 0)$ is equivalent to $(Eg) (x) (g(x, x + 17) = 0)$ which is equivalent to $(Eg) (x) (g(J(x, x + 17)) = 0)$ by (1). Likewise, $(Ex) (f) R(f, x)$ is equivalent to $(g) (Ex) R(\lambda_y g(x, y), x)$. (This paragraph assumes the axiom of choice.)

(3) Function quantifiers of like quality can be contracted to one if they are in immediate succession. This is based on the fact that $(Ef)(Eg) R(f, g)$ is equivalent to $(Ef) R(\lambda_y K(f(y)), \lambda_y L(f(y)))$, where K and L are the functions we introduced in the preceding section. E.g., $(Ef) (Eg) (x) (f(x) = g(x + 1))$ is equivalent to $(Ef) (x) (K(f(x)) = L(g(x + 1)))$. Likewise $(f)(g) R(f, g)$ is equivalent to $(f) R(\lambda_y K(f(y)), \lambda_y L(g(f(y))))$.

(4) A function quantifier and a number quantifier of like quality can be contracted to just a function quantifier, if they are in immediate succession. This is based on the fact that $(Ef) (Ex) R(f, x)$ is equivalent to $(Ef) R(\lambda_y K(f(y)), L(f(0)))$. E.g., $(Ef) (Ex) (f(x) = x + 3)$ is equivalent to $(Ef) (K(f(L(f(0)))) = L(f(0)) + 3)$. Likewise, $(f)(x) R(f, x)$ is equivalent to $(f) R(\lambda_y K(f(y)), L(f(0)))$.

(5) Number quantifiers can be reduced to one by using a function quantifier. This is based on the fact that, for example, $(x) (Ey) (z) (Ew) R(x, y, z, w)$ is equivalent to $(Ef) (Eg) (x) (z) R(x, f(x), z, g(x, z))$. Contracting the number quantifiers and contracting the function quantifiers we obtain the equivalent $(Ef)(x) R(K(x), K(f(K(x))), L(x), L(f(J(K(x), L(x))))$. Alternatively, we could first have taken the negation of $(x) (Ey) (z) (Ew) R(x, y, z, w)$, which is $(Ex) (y) (Ez) (w) - R(x, y, z, w)$. Applying the preceding technique to this negation, we obtain $(Ex) (Ef) (y) (w) - R(x, y, f(x), w)$. Negating again, to get something equivalent to the original formula, we get $(x) (f) (Ey) (Ew) R(x, y, f(x), w)$. Contracting the two universal quantifiers and the two existential quantifiers, we obtain the equivalent $(f) (Ey) R(L(f)(0), K(y), K(f(L(f)(0))), L(y))$. Thus an arithmetical formula can be rewritten with a prefix of either of the two forms $(Ef) (x)$ or $(f) (Ex)$.

(6) If both number quantifiers and function quantifiers are present, the number quantifiers can be reduced to one without increasing the number of function quantifiers. *Proof:* the function quantifiers can be advanced ahead of all the number quantifiers by (2). The predicate that follows all the function quantifiers now begins with a string of number quantifiers. If the last function quantifier is existential (universal) rewrite this predicate in the form $(Ef) (x)$ (respectively $(f) (Ex)$) by the methods of (5). Then contract the two successive existential (universal) quantifiers.

We can sum up (1) through (6) by saying that any analytical predicate can be written in a form in which the prefix consists of function quantifiers followed by a single number quantifier, and in which existential and universal quantifiers strictly alternate. If there are n function quantifiers in this form and the first is existential (universal) then the predicate is called a Σ_n^1 (respectively Π_n^1) predicate; what we previously called Σ_n and Π_n predicates, are now called Σ_n^0 and Π_n^0 predicates, when confusion with analytical predicates is possible. The superscript (0 or 1) denotes the *type* of the quantifiers of highest type, in the sense of roughly Russell's theory of types (taking the natural numbers as type 0, however, which is very un-Russellian); the subscript n denotes the number of alternating quantifiers of highest type.

The hierarchy theorem for the arithmetical hierarchy asserts that for every n , there is a Σ_{n+1}^0 set which is not a Π_{n+1}^0 set (and hence not a Σ_n^0 or a Π_n^0 set), and also a Π_{n+1}^0 set which is not a Σ_{n+1}^0 set. The analogous hierarchy theorem for the analytic hierarchy, whose proof will not be reviewed here, asserts that for every n , there is a Σ_{n+1}^1 set which is not a Π_{n+1}^1 set (and hence not a Π_n^1 or a Σ_n^1 set), and also a Π_{n+1}^1 set which is not a Σ_{n+1}^1 set.

8. Extending the arithmetical hierarchy. As we saw, the Σ_n^0 sets are just the sets that bear the n th power of the relation "r.e. in" to recursive sets, and hence just the sets many-one reducible to K^n . This gives us a very neat characterization of the effect of **number quantifiers**: an existential (universal) number quantifier just takes us from a predicate to another predicate (respectively the complement of a predicate) which is r.e. in the first predicate. Also, if a class of predicates contains the recursive

predicates and is closed under truth functions and bounded number quantifiers (quantifiers of the forms “for all x less than y ” and “there exists an x less than y ”), then every predicate which is r.e. in a predicate of the class is the existential quantification of a member of the class. This suggests looking for a similar characterization of the effect of **function quantifiers**.

To put it crudely: an existential quantifier “means” *r.e. in*. What does a function quantifier “mean”? This question has led to years of research on the part of many investigators.

Since the sets of numbers we can define if we add function quantifiers to first order arithmetic are just the analytic sets of numbers, this problem is just the problem of a representation theorem for the analytical hierarchy. We have just said what the Σ_n^0 sets are; *what are the Σ_n^1 sets?*

The obvious way to try to answer these questions is to try to extend the arithmetical hierarchy. If we can extend it in some natural way so that the extended hierarchy contains all analytic sets of numbers, then doubtless we will be able to “look” at the extension and “see” what adding a function quantifier does to the place of a set or predicate in the hierarchy.

But how should we go about extending the arithmetical hierarchy? In a sense, the entire arithmetical hierarchy is encapsulated in the following ω -sequence of sets: K, K', K'', K''', \dots . This suggests the first move to make: try extending this sequence of “complete” sets into the transfinite.

The classical approach (due independently to Davis and Mostowski) was to first encode the ordinals (up to some fixed countable ordinal) by integers. Let us use “ N_α ” to denote the set of integer “notations” for the ordinal α . (This assumes some fixed encoding has been adopted.) Then we may now take:

DEFINITION I.

$$K_0 = \emptyset.$$

$$K_{\alpha+1} = (K_\alpha)'. \quad (\text{Note that } K_1 = K.)$$

$$K_\lambda = \{J(m, n) \mid (Ex) (\alpha < \lambda \ \& \ n \in N_\alpha \ \& \ m \in K_\alpha)\},$$

(λ is a special variable for limit ordinals).

What this says is that at limit ordinals λ the corresponding “complete set” K_λ is a kind of recursive union of the previous K_α . This inductive definition associates a set K_α with each ordinal α for which there is a “notation” in the selected “encoding”. Davis and Mostowski used a more complicated definition of the sets, leading not to a unique set K_α associated with each ordinal α for which they had a notation, but to a set H_n associated with each *notation* n . It then had to be proved (as it was later, by Spector) that if n, m belong to the same N_α , then H_n is Turing equivalent to H_m (for the particular system of encoding employed by these authors, the system “ $<_0$ ” of Church and Kleene). When this was proved by Spector, it was possible to see that a

genuine hierarchy of degrees of unsolvability associated with ordinals had been defined, even though in the definition *sets*, not *degrees*, were associated with *notations*, not ordinals. However, it has been pointed out by Luckham and myself, and also independently by Enderton, that the above definition, leading to a unique set K_α associated with each ordinal, is essentially equivalent to the Davis-Mostowski procedure.

The set K_α , unfortunately, depends not only on the ordinal α but on the *notation system* (the “system of encoding”) employed. How bad is this?

The answer is: “very bad”. One can get as high a degree of unsolvability as one likes already for K_ω by choosing a suitable notation system. Since our aim is to extend the sequence K', K'', K''', \dots in some invariant way by defining sets $K_\omega, K_{\omega+1}, K_{\omega+2}, \dots$ etc., and *notations* for ordinals are only to be an auxiliary in this task, we find this extremely unsatisfactory.

Through the constructive ordinals a resolution of this difficulty is available. Each constructive ordinal is the ordinal of a recursive well ordering. So a natural “notation system” for the ordinals less than α , where α is any constructive ordinal, is *any recursive well ordering of order type α* . If R is a well ordering, or even a well founded partial ordering, we define the natural map from $\text{Field}(R)$ onto the ordinals in the appropriate segment as follows:

- (1) If $n \in \text{Field}(R)$ is R -minimal, then $|n|^R = 0$ (“ n is an R -notation for 0”).
- (2) If $m, n \in \text{Field}(R)$ and m is an immediate R -successor of n , then $|m|^R = |n|^R + 1$.
- (3) If $m \in \text{Field}(R)$ is an R -limit, then $|m|^R$ = the least upper bound of $\{|n|^R \mid Rnm\}$.

It turns out that if we regard any recursive well ordering as a notation system in this way, and choose that notation system as the system we use in the above definition of the sets K_α (writing $|n|^R = \alpha$ for $n \in N_\alpha$), then the *degrees* of the resulting sets K_α are independent of α . This theorem (which might be called the **external uniqueness theorem** for recursive well orderings) was proved by Enderton and Luckham. The same theorem holds for r.e. well orderings; and this result has the Spector uniqueness theorem for the system $<_0$ as a special case.

(The Spector uniqueness theorem is that if n, m are two notations for the same ordinal in the Church-Kleene system $<_0$, then the sets H_n, H_m are of the same degree. The definition of the sets H_n is more complicated than that given above for the sets K_α ; but if one takes the notations m in the system $<_0$ which bear the partial ordering relation $<_0$ to a given notation, then (1) the resulting “path” in the system $<_0$ is a recursively enumerable well ordering; and (2) the sets H_n along the path are of the same Turing degrees as the sets K_α *along the path*—i.e., thinking of the path as a system of notations, and ignoring the rest of $<_0$.)

The result of Enderton and Luckham justifies the following definition for constructive ordinals α (the “constructive ordinals” were originally defined in terms of

the system $<_0$ —in fact, as the ordinals for which there are notations in that system, but it is convenient to redefine them simply as the ordinals of recursive well orderings):

d_α = the degree of unsolvability of K_α , where K_α is the set associated with the ordinal α by Definition I if we use an arbitrary recursive system of notations (i.e., recursive well ordering or well founded partial ordering) which contains a notation for α .

Thus a true hierarchy of degrees of unsolvability has been associated with constructive ordinals in a satisfactorily invariant manner.

9. Extending beyond the constructive ordinals. The sets K_α , α a constructive ordinal, and the sets recursive in them, together form only the *hyperarithmetical* sets (assuming we use recursive systems of notations, as just proposed), and these are only the tiniest fraction of the analytic sets—in fact, just the sets in $\Sigma_1^1 \cap \Pi_1^1$. Our original program was to extend the sequence K, K', K'', \dots until we had (at least) all the *analytic* sets. We have certainly not succeeded in this! At present various attempts have been made to extend farther. The farthest extension yet proposed is the hierarchy of “constructible degrees of unsolvability” proposed by Boolos and the author [2]; to show that this extension, which it goes beyond the scope of the present paper to discuss, includes all sets of integers or even all analytic sets of integers, requires the assumption of the controversial axiom “ $V = L$ ” suggested by Gödel (but not longer advocated by him) as a new axiom for general set theory in his paper on the consistency of the axiom of choice and the generalized continuum hypothesis. In the absence of “ $V = L$ ”, or some such axiom it remains consistent that the extension referred to does not include all analytic sets of integers. Moreover, even if we assume “ $V = L$ ”, no good representation theorem for the analytic hierarchy has yet been proved by anyone. The preceding section of this paper is the barest introduction to a fascinating and totally unsolved problem; the problem of obtaining an understanding of the analytic hierarchy from the viewpoint of recursive function theory.

References

1. Hartley Rogers, Jr., *Theory of Recursive Functions and Effective Computability*, McGraw-Hill, New York, 1967. This work contains an excellent bibliography of the subject.
2. George Boolos and Hilary Putnam, Degrees of unsolvability of constructible sets of integers, *J. Symbolic Logic*, vol. 33, no. 4 (December 1968), pp. 497–513.

FUNCTION THEORY ON SOME NONARCHIMEDEAN FIELDS

ABRAHAM ROBINSON, Yale University

1. Introduction. Archimedes' axiom states that for any two positive numbers a and b , a smaller than b , the continued addition of a to itself ultimately yields numbers which are greater than b . More formally, if F is an ordered abelian group or, more particularly, an ordered field, then Archimedes' axiom is as follows.

1.1. *If $0 < a < b$, where a and b are elements of F then there exists a natural number n such that*

$$\underbrace{a + a + \cdots + a}_{n \text{ times}} > b.$$

Throughout the history of mathematics, Archimedes' axiom has been associated with the foundations of the Differential and Integral Calculus. Already in Greek science the method which, much later, was dubbed the method of exhaustion and which, to a large extent, anticipated the ϵ, δ method in the calculation of areas and volumes, depended on the validity of Archimedes' axiom, which was formulated explicitly for this purpose. On the other hand, when a method of infinitely small and infinitely large numbers is used, as in Nonstandard Analysis, then it is just the non-archimedean nature of the system which is essential for its success or, more precisely, the superposition of a nonarchimedean field on the archimedean field of real numbers.

Although Nonstandard Analysis (see [4] or [6]) may perhaps be regarded as the most successful effort in this direction, many other systems have been introduced for the same purpose. Thus, not long ago, D. Laugwitz [2] considered a theory of functions on the field L of generalized power series with real coefficients and **real exponents**. The same field was investigated many years earlier by T. Levi-Civita [3], also because of its nonarchimedean character, and by A. Ostrowski [5], in connection with the theory of valuations.

Laugwitz raised the question whether the functions considered by him satisfy the intermediate value theorem and the mean value theorem of the Differential Calculus. We shall show in the present paper that although these theorems are not valid here in full generality, they are true under rather wide conditions. In order to obtain these results, we shall embed L in the residue class field oR of a certain subring of a nonstandard model of Analysis, *R . It appears that oR has many interesting properties which make it a suitable subject for investigation quite apart from the particular problem just mentioned. In particular, the behavior of a function on oR is closely connected with the theory of asymptotic expansions, although we shall not pursue this topic in the present paper.

2. Ordered fields and fields with valuation. An ordered field F is a commutative field in which an ordering relation $x < y$ (or, equivalently, $y > x$) is defined and satisfies the following conditions.

2.1. *The ordering is transitive, $x < y$ and $y < z$ implies $x < z$, and irreflexive, $x < y$ implies $x \neq y$.*

2.2. *The ordering is total, if $x \neq y$ then either $x < y$ or $y < x$ (but not both, by 2.1).*

2.3. *The ordering is related to addition by the requirement that $x < y$ implies $x + z < y + z$; and to multiplication by the requirement that $x < y$ and $0 < z$ implies $xz < yz$.*

An ordered field can be characterized also by means of the set of its **positive elements** $P = \{x \mid x > 0\}$. Thus, suppose that a subset P of a field F possesses the following properties.

2.4. $0 \notin P$; for all $x \neq 0$, $x \in P$ or $-x \in P$.

2.5. If $x, y \in P$, then $x + y \in P$ and $xy \in P$.

Then the relation *defined* by

$$x < y \text{ if and only if } y - x \in P$$

satisfies the conditions 2.1–2.3 and P is just the set of positive elements of the field according to this relation.

We shall suppose that the reader is familiar with the elementary properties of ordered fields, e.g., that an ordered field is of characteristic 0 and that $x^2 > 0$ for all $x \neq 0$. As usual, we write $x \leq y$ or $y \geq x$ if either $x < y$ or $x = y$.

The rational numbers form an ordered field Q whose positive elements are the fractions (ratios) of natural numbers different from zero, and the real numbers form an ordered field R whose positive elements are just the squares other than zero. In both cases the ordering is unique. Moreover, both Q and R are archimedean, i.e., they satisfy Archimedes' Axiom 1.1.

Perhaps the simplest example of a non-archimedean field is as follows. Let $R(t)$ be a simple transcendental extension of the field of real numbers R . Thus $R(t)$ may be identified with the field of rational functions of the indeterminate t with coefficients in R , each element of $R(t)$ may be written in the form

$$(2.6) \quad f = \frac{p(t)}{q(t)} = \frac{a_0 + a_1 t + \cdots + a_n t^n}{b_0 + b_1 t + \cdots + b_m t^m},$$

where $q(t) \neq 0$, at least one of the b_j is different from 0. We may then suppose the first $b_j \neq 0$ is actually equal to 1, for if this is not the case from the outset, we may achieve it by multiplying the numerator and denominator on the right hand side of (2.6) by b_j^{-1} . Thus, if $f \neq 0$, we may write

$$(2.7) \quad f = \frac{a_k t^k + \cdots + a_n t^n}{t^j + b_{j+1} t^{j+1} + \cdots + b_m t^m}, \quad a_k \neq 0,$$

$$0 \leq k \leq n, \quad 0 \leq j \leq m.$$

We now determine an ordering in $R(t)$ by defining that $f \neq 0$ is **positive** if and only if $a_k > 0$. To make sure that this is a good definition one first has to check that it is independent of the particular representation (2.7) chosen for the given f . Next one verifies that the set of positive elements of $R(t)$ defined in this way satisfies the conditions of 2.4 and 2.5. We suppose that these rather simple tasks have been carried out so that $R(t)$ becomes indeed an ordered field with the above definition. Moreover, this ordered field is nonarchimedean. For, by our definitions, $0 < t, t < 1$ (since $1 - t$ is positive) and, for any positive integer n ,

$$\frac{t + t + \cdots + t}{n \text{ times}} < 1$$

(since $1 - nt$ is positive). This shows that 1.1 is not satisfied.

In any ordered field, the absolute value of a number a is defined to be $|a| = a$ if $a \geq 0$, otherwise $|a| = -a$. Then $|ab| = |a| |b|$ and $|a + b| \leq |a| + |b|$ (triangle inequality).

Let F be a nonarchimedean ordered field. Then F is of characteristic 0 and, hence, contains the field of rational numbers Q . An element $a \in F$ is said to be **infinite** if $|a| > q$ for all $q \in Q$. Also, $a \in F$ is said to be **infinitely small** or **infinitesimal** if $|a| < q$ for all positive $q \in Q$. $a \in F$ is **finite** if it is not infinite. This will be the case if and only if $|a| < q$ for some $q \in Q$.

The finite elements of F constitute a subring F_0 of F . The infinitesimal elements of F constitute a proper ideal F_1 within F_0 . F_1 is maximal in F_0 as can be seen by the following argument. Suppose that $F_1 \subset J \subset F_0$ where J is an ideal in F_0 , such that $J - F_1 \neq \emptyset$. Let $a \in J - F_1$ then a is not infinitesimal. We conclude without difficulty that a^{-1} is finite, so $a^{-1} \in F_0$, $aa^{-1} = 1 \in J$. But then $J = F_0$, F_1 is maximal in F_0 .

It follows that $F' = F_0/F_1$ is a field. F' is called the **residue class field of the ordering**. The canonical mapping $F_0 \xrightarrow{\psi} F'$ induces an ordering in F' according to the rule that, for any $a \in F'$, $a \neq 0$, a is to be positive in F' if and only if one (and hence, all) of the elements of $\psi^{-1}a$ is (are) positive. It is not difficult to show that F' is archimedean according to this ordering and (hence) that it is isomorphic and order-isomorphic to a subfield of R .

The cosets of F_1 as an additive subgroup of F are called **monads**. If a is any element of F then we denote the monad containing it by $\mu(a)$. In particular, $\mu(0) = F_1$. The monads which are subsets of F_0 may be identified with the elements of F' .

As a tool in our investigation of nonarchimedean fields we shall require also the notion of a **field with valuation**, more particularly, the notion of a **field with non-**

archimedean valuation in the real numbers. This concept is given by a field F together with a mapping $v(x)$ from $F - \{0\}$ into the real numbers R such that the following conditions are satisfied:

2.8. For all $x \neq 0, y \neq 0$ in $F, v(xy) = v(x) + v(y)$.

2.9. For all x, y in F such that $x \neq 0, y \neq 0, x + y \neq 0$,

$$v(x + y) \geq \min(v(x), v(y)).$$

If we add to R an **element** ∞ (usually called "a symbol") with the rules $x + \infty = \infty + x = \infty + \infty = \infty$ and the stipulation that $\infty > x$ for all real x , then the auxiliary definition $v(0) = \infty$ ensures that the equations of 2.8 and 2.9 are satisfied without any restriction on x and y .

The set $O_F = \{x \in F \mid v(x) \geq 0\}$ is a subring of F , the **valuation ring**, and the set $J_F = \{x \in F \mid v(x) > 0\}$ constitutes a maximal ideal in O_F , the **valuation ideal**. The field $\bar{F} = O_F/J_F$ is called the residue class field of the given valuation.

Let c be an arbitrary but fixed constant greater than 1. Then the definition of distance

$$d(x, y) = c^{-v(x-y)},$$

where $c^{-\infty}$ is interpreted as 0, turns F into a metric space. If every Cauchy sequence in that space has a limit then F is said to be **complete** for the given valuation.

See [1], [7] or [8] for basic facts in valuation theory. From now on such facts will be taken for granted.

3. The field L . The field $R(t)$ is inadequate for the development of the calculus because we cannot extend to it even some of the most common functions defined in the field of real numbers, e.g., the function $y = \sqrt{x}$. Passing to the field of formal Laurent series $\sum_{k=-\infty}^{\infty} a_k t^k, a_k \in R$, does not remedy the situation. Following Laugwitz, we therefore consider the field of generalized power series L , which is defined as follows:

The elements of L are the formal expressions

$$(3.1) \quad \sum_{k=0}^{\infty} a_k t^{v_k} \quad a_k, v_k \in R, \quad v_k \uparrow \infty,$$

(where the last symbol implies $v_0 < v_1 < v_2 < \dots$). Two expressions (3.1) are, by definition regarded as equal if for any term a_v which occurs in one but not in the other, $a = 0$. We shall also write $a_0 t^{v_0} + a_1 t^{v_1} + \dots + a_k t^{v_k}$ for an expression for which $a_{k+1} = a_{k+2} = \dots = 0$.

The **sum** of two expressions $\sum a_k t^{v_k}$ and $\sum b_k t^{\mu_k}$ as in (3.1) is the expression $\sum c_k t^{\lambda_k}$ which is defined as follows. The sequence $\{\lambda_k\}$ is the set theoretical union of the sequences $\{v_k\}$ and $\{\mu_k\}$ arranged in increasing order. If a particular λ_m occurs both in $\{v_k\}$ and in $\{\mu_k\}$, e.g., $\lambda_m = v_p = \mu_q$ then $c_m = a_p + b_q$; if $\lambda_m = v_p$ but λ_m does not occur in $\{\mu_k\}$ then $c_m = a_p$; and if $\lambda_m = \mu_q$ but λ_m does not occur in $\{v_k\}$ then

$c_m = b_q$. Thus, briefly, the sum $\sum c_k t^{\lambda_k}$ is obtained by the formal addition of the terms of $\sum a_k t^{v_k}$ and $\sum b_k t^{\mu_k}$. Similarly, the **product** $\sum c_k t^{\lambda_k}$ of $\sum a_k t^{v_k}$ and $\sum b_k t^{\mu_k}$ as in 3.1 is obtained by formal multiplication. Thus, the sequence $\{\lambda_k\}$ consists of the sums $v_p + \mu_q$ arranged in increasing order and $c_k = \sum a_p b_q$ where p and q range over the natural numbers such that $v_p + \mu_q = \lambda_k$. It is not difficult to see that all these sums are finite and that the resulting expression satisfies the conditions of (3.1). Moreover, our definitions of sums and products are compatible with the relation of equality introduced earlier, and they turn L into a ring whose zero and unit elements may be written as $0t^0 + 0t^1 + 0t^2 + \dots$, or 0, and as $1t^0 + 0t^1 + 0t^2 + \dots$, or 1.

Now let $\alpha = 1 + \sum_{k=1}^{\infty} a_k t^{v_k}$, $0 < v_1 < v_2 < \dots \rightarrow \infty$, i.e., α is an element of L as in 3.1 with $v_0 = 0$, $a_0 = 1$. We wish to show that α possesses a multiplicative inverse in L . For this purpose we define β as the formal expansion in powers of t of the expression

$$1 - \left(\sum_{k=1}^{\infty} a_k t^{v_k} \right) + \left(\sum_{k=1}^{\infty} a_k t^{v_k} \right)^2 - \left(\sum_{k=1}^{\infty} a_k t^{v_k} \right)^3 + \dots$$

Again it is not difficult to see that this expansion can be worked out and that it is of the form $\beta = 1 + \sum_{k=1}^{\infty} b_k t^{\mu_k}$ where $0 < \mu_1 < \mu_2 < \dots \rightarrow \infty$, so that β belongs to L .

We now claim that $\alpha\beta = 1$. To see this, consider the identity

$$(3.2) \quad (1 + \gamma)(1 - \gamma + \gamma^2 - \gamma^3 + \dots + \gamma^{2m}) = 1 + \gamma^{2m+1}$$

which holds in L for arbitrary natural m . We may substitute $\sum_{k=1}^{\infty} a_k t^{v_k}$ for γ and expand on both sides of (3.2). This yields an equation

$$(3.3) \quad \alpha\beta' = \gamma',$$

where β' is the expansion of

$$1 - \left(\sum_{k=1}^{\infty} a_k t^{v_k} \right) + \left(\sum_{k=1}^{\infty} a_k t^{v_k} \right)^2 - \dots + \left(\sum_{k=1}^{\infty} a_k t^{v_k} \right)^{2m}$$

and γ' is the expansion of $1 + \left(\sum_{k=1}^{\infty} a_k t^{v_k} \right)^{2m+1}$. But then β' differs from β only in powers of t whose exponent is at least $(2m+1)v_1$ and γ' differs from 1 only in powers of t whose exponent also is at least $(2m+1)v_1$. Since m is an arbitrary natural number, we conclude that $\alpha\beta = 1$, $\beta = \alpha^{-1}$.

Now let $\alpha \in L$ be different from zero, otherwise arbitrary. Then $\alpha = \sum_{k=0}^{\infty} a_k t^{v_k}$, where we may assume that $a_0 \neq 0$. Putting $\alpha = a_0 t^{v_0} \alpha'$ where

$$\alpha' = 1 + \sum_{k=1}^{\infty} (a_k/a_0) t^{v_k - v_0},$$

we then obtain $a_0^{-1} t^{-v_0} \alpha'^{-1}$ as the multiplicative inverse of α .

Thus, L is a field. We introduce an ordering of L by defining that an element

$\alpha \in L$, $\alpha \neq 0$ is positive if and only if the nonvanishing coefficient a_k with lowest subscript m in the expression $\alpha = \sum_{k=0}^{\infty} a_k t^{v_k}$ is positive. Also, L obtains a valuation by defining $v(\alpha) = v_m$ (so that $a_m \neq 0$, $a_k = 0$ for $k < m$), for $\alpha \neq 0$, together with $v(0) = \infty$ in accordance with our general convention.

In this valuation, the valuation ring O_L consists of all elements of L which can be written as $\sum a_k t^{v_k}$ with $v_0 \geq 0$, and this is also the ring of **finite** elements of L in the ordering of L ; and the valuation ideal J_L consists of all $\sum a_k t^{v_k}$ with $v_0 > 0$ and coincides with the set of **infinitesimal** elements of L . Thus, the residue class field of L with respect to its valuation coincides with the residue class field of L with respect to its ordering and is, in fact, the field of real numbers R . Also, since $J_L \neq \{0\}$, L is nonarchimedean.

There is a natural (and obvious) embedding (injection) of R into L : $a \rightarrow a = at^0 + 0t^1 + 0t^2 + \dots$ and this extends, equally obviously, to an embedding of $R[t]$ into L :

$$a_0 + a_1 t + \dots + a_n t^n \rightarrow a_0 t^0 + a_1 t^1 + a_2 t^2 + \dots + a_n t^n + 0t^{n+1} + \dots$$

and hence, to an embedding of $R(t)$ into L . The embedding is order preserving for the ordering of $R(t)$ defined in section 2 above.

It is shown in [5] that L is complete. It is also shown there that the field L' which is obtained by taking complex coefficients in place of the real coefficients in L , is algebraically closed. Since $L' = L(\sqrt{-1})$ it follows (compare [7]) that L is real-closed, i.e., that every positive element of L possesses a square root in L and that every polynomial of odd degree in $L[x]$ possesses a root in L . It follows in particular that a positive element of L possesses roots of all orders $n = 2, 3, 4, \dots$. The same result is established by elementary means in [2] and will be used later in this paper.

Now let $f(x)$ be a real-valued infinitely differentiable function of a real variable which is defined in an interval $a < x < b$, $a, b \in R$. On passing from R to L , we find that the interval $a < x < b$ in L consists of points $x = \xi + \sum_{k=1}^{\infty} a_k t^{v_k}$, $0 < v_1 < \dots \rightarrow \infty$, of three kinds,

- (i) $a < \xi < b$,
- (ii) $\xi = a, \sum_{k=1}^{\infty} a_k t^{v_k} > 0$, and
- (iii) $\xi = b, \sum_{k=1}^{\infty} a_k t^{v_k} < 0$.

In all these cases ξ is the unique real number which is infinitely close to x , i.e., such that $x - \xi$ is infinitely small and (by analogy with the terminology in Nonstandard Analysis) we call ξ the **standard part** of x , $\xi = {}^0x$.

Laugwitz extends the function $f(x)$ to values of x in L with standard part ξ , $a < \xi < b$ by using the formal Taylor expansion of $f(x)$,

$$f(x+h) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(x) h^n.$$

Thus, he defines for $x = \xi + \sum_{k=1}^{\infty} a_k t^{v_k}$,

$$(3.4) \quad {}^L f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(\xi) \left(\sum_{k=1}^{\infty} a_k t^{v_k} \right)^n,$$

where it is understood that ${}^L f(x)$ is the element of L which is obtained by expanding the right hand side of (3.4) and rearranging it in powers of t . Once again, the condition $v_0 > 0$ shows that this can be done.

We shall show in the following sections that the definition proposed by Laugwitz is obtained in a natural way by relating L to a nonstandard model of Analysis.

4. The field *R . Let *R be a nonstandard model of Analysis (cf. [4] and [6]). We shall suppose that *R is **sequentially comprehensive**. That is to say, if $a_0, a_1, a_2, \dots, a_n, \dots, n \in N$, is a sequence of entities of *R (of the same type, if type restrictions are adopted), e.g., a sequence of numbers of *R , then there exists an internal sequence $\{s_n\}$ in *R (where n now ranges over *N) such that $s_n = a_n$ for all finite n .

*There exist sequentially comprehensive *R .* More particularly, all *R which are ultrapowers are sequentially comprehensive. Thus, suppose ${}^*R = R^I/D$ where D is a free ultrafilter on the index set I . Every internal entity of *R is represented by (is an equivalence class of) functions $f(v)$ on I . Let $f_n(v)$ represent a_n , $n = 0, 1, 2, \dots$, and for each $v \in I$, consider $s(v) = \{f_n(v)\}$. Then $s(v)$, v ranging over I represents an internal sequence $\{s_n\}$ in *R . We claim that for each finite k , the value of that sequence is just a_k . Now, in order to obtain the value of $\{s_n\}$ for $n = k$, we have to substitute the function $f(v) \equiv k$ for each n in $f_n(v)$. This yields precisely $f_k(v)$, i.e., a_k .

Supposing, from now on, that *R is sequentially comprehensive, we wish to show that the set of infinite natural numbers, ${}^*N - N$, cannot be coinital with ω^* . In other words:

4.1. THEOREM. *Let $a_0 > a_1 > a_2 > \dots > a_n > \dots, n \in N$ be a strictly decreasing sequence of infinite natural numbers, internal or external. Then there exists an infinite natural number a , such that $a_n > a$ for all $n \in N$.*

Proof. Since *R is sequentially comprehensive, we may suppose that, for all $n \in N$, $a_n = s_n$ where $\{s_n\}$ is an internal sequence of numbers of *R . Consider the internal sequence

$$t_n = \frac{n}{\min(s_0, s_1, \dots, s_n)}, \quad n \in {}^*N.$$

Then $0 \leq t_n < 1$ for all finite n but $t_n > 1$ for large infinite n . Hence there exists a smallest m , which must then be infinite, such that $0 \leq t_m < 1$ does not hold.

Thus, for $k = m - 1$,

$$0 \leq \frac{k}{\min(s_0, s_1, \dots, s_k)} < 1.$$

This shows that $k < a_0, k < a_1, \dots, k < a_n, \dots$ for all finite n and proves the theorem.

Now let ρ be an arbitrary but fixed positive infinitesimal number in *R . We define subsets M_0 and M_1 of *R by

$$M_0 = \{x \in {}^*R \mid |x| \leq \rho^{-n} \text{ for some finite positive integer } n\},$$

$$M_1 = \{x \in {}^*R \mid |x| \leq \rho^n \text{ for all finite positive integers } n\}.$$

Evidently, $M_1 \subset M_0$ and $M_0 \supset R$. Both M_0 and M_1 are rings under the operations of *R . For if $|x| \leq \rho^{-n}, |y| \leq \rho^{-m}$, with $n \leq m$ say, then

$$|x + y| \leq |x| + |y| \leq 2\rho^{-m} \leq \rho^{-(m+1)}$$

and $|xy| \leq \rho^{-(n+m)}$, so M_0 is a ring. And if $|x| \leq \rho^n, |y| \leq \rho^n$ then $|x \pm y| \leq 2\rho^n \leq \rho^{n-1}, |xy| \leq \rho^{2n}$. Since, in the definition of M_1 , n is arbitrary, this shows that M_1 also is a ring.

Moreover, M_1 is an ideal in M_0 , for if $x \in M_1$ and $y \in M_0$ then $|y| \leq \rho^{-n}$ for some natural number n , and since $|x| \leq \rho^{m+n}$ for all natural n , it follows that $|xy| \leq \rho^m$ for all natural m , $xy \in M_1$. M_1 is a proper ideal since it does not contain 1. Finally, M_1 is a *maximal* ideal in M_0 . For let $J \supset M_1$ be another ideal in M_0 such that $J - M_1$ is not empty, and let $x \in J - M_1$. Then $|x| > \rho^m$ for some finite natural number m and so $|x^{-1}| < \rho^{-m}$, $x^{-1} \in M_0$. Hence $1 = xx^{-1} \in J$, $J = M_0$, showing that M_1 is maximal.

We conclude that the quotient ring ${}^oR = M_0/M_1$ is a field. Moreover, the canonical map

$$(4.2) \quad \psi: M_0 \rightarrow {}^oR$$

induces an ordering in oR . For let $x \in M_0 - M_1$, $x > 0$, and let $x + y, y \in M_1$ be any other element of the coset of x with respect to M_1 . Then $|x| > \rho^m$ for some finite natural number m and $|y| \leq \rho^n$ for all finite natural numbers n . Hence $|y| < |x|$, and so $x + y \geq x - |y| = |x| - |y| > 0$, all elements of the coset of x are positive. Accordingly, we may define an ordering in oR by defining that an element $\alpha \in {}^oR$, $\alpha \neq 0$, is positive if and only if the elements of $\psi^{-1}\alpha$ are positive. Then the sum and product of positive elements of oR are positive but $0 \in {}^oR$ is not positive. This shows that our definition turns oR into an ordered field. We also observe that for any $\alpha \in {}^oR$, $\psi^{-1}\alpha$ is an interval in M_0 and *R . Finally, since M_1 contains only the single standard number 0, ψR provides an embedding of R (as a subfield of *R) in oR .

Next, we define a valuation in oR , as follows. For any $\alpha \in {}^oR$, $\alpha \neq 0$, let x and $x + y$ be elements of $\psi^{-1}\alpha$, $y \in M_1$ and consider $\log_\rho |x|$ and $\log_\rho |x + y|$. Since $|x|$

and $|x + y|$ are greater than some positive, and smaller than some negative power of ρ , $\log_\rho |x|$ and $\log_\rho |x + y|$ are finite and possess standard parts. We claim that

$${}^o(\log_\rho |x|) = {}^o(\log_\rho |x + y|),$$

i.e., that

$$\log_\rho |x + y| - \log_\rho |x| = \log_\rho |1 + y/x|$$

is infinitesimal. But $\log_\rho |1 + (y/x)| = \ln |1 + (y/x)| / \ln \rho$. Since y/x is infinitesimal and $\ln |w|$ is a standard function which is continuous at $w = 1$, $\ln |1 + (y/x)|$ is infinitesimal. Hence $\log_\rho |1 + (y/x)|$ also is infinitesimal, as asserted.

Accordingly, we obtain a unique definition of a function $v(\alpha)$ for $\alpha \in {}^oR$, $\alpha \neq 0$, by putting $v(\alpha) = {}^o(\log_\rho |x|)$ for any $x \in \psi^{-1}\alpha$. We claim that this defines a valuation of the field oR .

Let $\alpha, \beta \in {}^oR$, $\alpha \neq 0$, $\beta \neq 0$ and let $x \in \psi^{-1}\alpha$, $y \in \psi^{-1}\beta$. Then

$${}^o(\log_\rho |xy|) = {}^o(\log_\rho |x|) + {}^o(\log_\rho |y|)$$

and so $v(\alpha\beta) = v(\alpha) + v(\beta)$, as required. Next, suppose $\alpha + \beta \neq 0$, then we have to show that $v(\alpha + \beta) \geq \min(v(\alpha), v(\beta))$ or, equivalently, that

$$(4.3) \quad {}^o(\log_\rho |x + y|) \geq \min({}^o(\log_\rho |x|), {}^o(\log_\rho |y|)).$$

We may suppose without essential loss of generality that $\log_\rho |x| \geq \log_\rho |y|$. Then (4.3) will hold precisely if there is an infinitesimal η such that

$$\log_\rho |x + y| \geq \log_\rho |y| - \eta,$$

i.e., such that

$$\log_\rho \left| 1 + \frac{x}{y} \right| \geq -\eta.$$

Putting $x/y = w$, we have to show $\log_\rho |1 + w| \geq -\eta$ for $\log_\rho |w| \geq 0$, (where we may rule out $w = -1$ because of $\alpha + \beta \neq 0$). Put $\sigma = \log_\rho |w|$, $|w| = \rho^\sigma$, where $\sigma \geq 0$, then

$$\begin{aligned} |1 + w| &\leq 1 + |w| = 1 + \rho^\sigma \leq 2\rho^\sigma = \rho^{\sigma + \log_\rho 2} \\ \log_\rho |1 + w| &\geq \sigma + \log_\rho 2 \geq \log_\rho 2. \end{aligned}$$

But $\ln_\rho 2$ is (negative) infinitesimal, and so (4.3) is proved. We supplement the definition of $v(x)$ as usual by putting $v(0) = \infty$.

The valuation ring of the valuation just defined will be denoted by O_ρ . It is not difficult to see that O_ρ includes the ψ -images of all finite elements of *R . However, O_ρ includes other elements as well. For example, let $\lambda = \psi \ln \rho$. Then $v(\lambda) = {}^o(\log_\rho |\ln \rho|) = {}^o(\ln |\ln \rho| / \ln \rho)$. But the expression in the parentheses on the right hand side is

infinitesimal, for $\ln \rho$ is (negative) infinite and

$$\lim_{x \rightarrow \infty} \frac{\ln x}{x} = 0.$$

Hence $v(\lambda) = 0$.

We shall now show that the field oR is *complete* for the valuation defined above. Defining the distance between two elements of oR , α and β , by $d(\alpha, \beta) = c^{-v(\alpha-\beta)}$ (see the end of section 2 above) let $\{\alpha_n\}$ be a Cauchy sequence in this metric.

$$(4.4) \quad \lim_{\substack{n \rightarrow \infty \\ m \rightarrow \infty}} d(\alpha_n, \alpha_m) = 0.$$

Then we have to show that $\{\alpha_n\}$ converges to a limit α in oR .

Choose elements $x_n \in \psi^{-1}\alpha_n$, $n = 0, 1, 2, \dots, n \in N$. Since *R is sequentially comprehensive there exists an internal sequence $\{s_n\}$ of elements of *R such that $s_n = x_n$ for all finite n . We shall write x_n in place of s_n also for infinite n . By (4.4)

$$\lim_{\substack{n \rightarrow \infty \\ m \rightarrow \infty}} v(\alpha_n - \alpha_m) = \infty.$$

Equivalently, given any finite natural number k , there exists a finite natural $j = j_k$ such that

$$(4.5) \quad \log_\rho |x_n - x_m| > k \text{ for } n, m > j_k, \quad n, m \in N.$$

Now since 4.5 holds for all finite n and m greater than j , it holds for all $n > j$, $m > j$, $n + m$ finite, $j = j_k$. A standard argument of Nonstandard Analysis, which was exemplified in the proof of 4.1, now shows that there exists an *infinite* natural $\omega = \omega_k$ such that (4.5) holds for all $n > j$, $m > j$, and $n + m < 2\omega_k$ and hence, in particular, for all $n > j$, $m > j$ and $n < \omega_k$, $m < \omega_k$. Moreover, by determining $\omega_0, \omega_1, \omega_2, \dots$ one after the other, we may evidently assume that $\omega_0 > \omega_1 > \omega_2 > \dots$. Appealing to 4.1, we may then choose an infinite natural number Ω which is smaller than $\omega_0, \omega_1, \omega_2$ and—obviously, being infinite, larger than j_0, j_1, j_2, \dots . Then,

$$(4.6) \quad \log_\rho |x_n - x_\Omega| > k \text{ for } n > j_k, \quad n \in N, \quad k \in N.$$

(4.6) shows in the first place, that $x_\Omega \in M_0$. To see this, choose $n > j_0$ then $\log_\rho |x_n - x_\Omega| > 0$, so $|x_n - x_\Omega|$ is finite. Also, $x_n \in M_0$, so $|x_n| \leq \rho^{-m}$ for some positive integer m and $|x_\Omega| \leq |x_\Omega - x_n| + |x_n| \leq 2\rho^{-m} < \rho^{-(m+1)}$, $x_\Omega \in M_0$.

Now let $\alpha = \psi x_\Omega$, then we wish to show that $\lim_{n \rightarrow \infty} \alpha_n = \alpha$ or, which is equivalent, that

$$(4.7) \quad \lim_{n \rightarrow \infty} v(\alpha_n - \alpha) = \infty.$$

But this is an immediate consequence of (4.6), since (4.6) implies

$$^0(\log_p |x_n - x_\Omega|) > k - 1 \text{ for } n > j_k, \quad n \in N, \quad k \in N$$

and this is the same as

$$v(\alpha_n - \alpha) > k - 1 \text{ for } n > j_k, \quad n \in N, \quad k \in N$$

which is just an explicit expression for the validity of (4.7). Thus, we have shown that 0R is complete.

Let $\bar{\rho} = \psi\rho$ and consider any infinite series in 0R of the form

$$(4.8) \quad a_0\bar{\rho}^{v_0} + a_1\bar{\rho}^{v_1} + a_2\bar{\rho}^{v_2} + \dots, \quad a_n \in R \subset {}^0R, \\ v_0 < v_1 < v_2 < \dots \rightarrow \infty,$$

where the v_j are standard real. The partial sums of (4.8) are

$$\sigma_k = a_0\bar{\rho}^{v_0} + a_1\bar{\rho}^{v_1} + \dots + a_k\bar{\rho}^{v_k}, \quad k = 0, 1, 2, \dots$$

The value of any monomial in (4.8) is, for $a_j \neq 0$, $v(a_j\bar{\rho}^{v_j}) = v(a_j) + v(\bar{\rho}^{v_j}) = 0 + v_j = v_j$, with $v(a_j\bar{\rho}^{v_j}) = \infty$ for $a_j = 0$. Hence $v(\sigma_k) = v_j$ where j is the smallest subscript $\leq k$ for which $a_j \neq 0$, if any, otherwise $v(\sigma_k) = \infty$. Also, for $0 \leq k < l$

$$\sigma_l - \sigma_k = a_{k+1}\bar{\rho}^{v_{k+1}} + \dots + a_l\bar{\rho}^{v_l}$$

and so $v(\sigma_l - \sigma_k) \geq v_{k+1}$. This shows that $\{\sigma_k\}$ is a Cauchy sequence, and the limit of that sequence, σ is just the sum of (4.8). Also, $v(\sigma) = v_j$ where j is the lowest subscript for which $a_j \neq 0$ or, if there is no such j , i.e., if all a_j vanish, $v(\sigma) = \infty$ and this is the case if and only if $\sigma = 0$. As usual in the theory of infinite series, we identify (4.8) with its sum in 0R . It is then not difficult to verify that the sum of two numbers of 0R , σ and τ , given by (4.8) and

$$(4.9) \quad b_0\bar{\rho}^{\mu_0} + b_1\bar{\rho}^{\mu_1} + b_2\bar{\rho}^{\mu_2} + \dots, \quad b_n \in R \subset {}^0R, \\ \mu_0 < \mu_1 < \mu_2 < \dots \rightarrow \infty,$$

is represented by an expression

$$c_0\bar{\rho}^{\lambda_0} + c_1\bar{\rho}^{\lambda_1} + c_2\bar{\rho}^{\lambda_2} + \dots$$

which is obtained from (4.8) and (4.9) just as the sum $\sum c_k t^{\lambda_k}$ was obtained from $\sum a_k t^{v_k}$ and $\sum b_k t^{\mu_k}$ as elements of L in section 3 above. The product of (4.8) and (4.9) also is obtained by the procedure described in section 3, with $\bar{\rho}$ for t . It follows that the mapping

$$(4.10) \quad \Phi: a_0 t^{v_0} + a_1 t^{v_1} + a_2 t^{v_2} + \dots \rightarrow a_0 \bar{\rho}^{v_0} + a_1 \bar{\rho}^{v_1} + a_2 \bar{\rho}^{v_2} + \dots, \\ a_j \in R, \quad v_0 < v_1 < v_2 < \dots \rightarrow \infty,$$

where the v_j are standard real, is a homomorphism from L into 0R . This homomorphism is an injection since $\Phi\alpha = 0$ implies $a_0 = a_1 = a_2 = \dots = 0$ (see above)

and, hence, $\alpha = 0$. It follows that ΦL is a field which is isomorphic to L and we write $\Phi L = {}^o L$. Evidently, Φ is analytic (i.e., value preserving, $v(\Phi(\alpha)) = v(\alpha)$). But Φ is also order preserving, as can be shown by verifying that, for any $\alpha \in L$, $\Phi\alpha > 0$ if and only if $\alpha > 0$. Now for $\alpha \neq 0$, $\alpha > 0$ if and only if the first nonvanishing a_j is positive, so we only have to show that an expression as in (4.8), $\bar{\sigma} = a_0\bar{\rho}^{v_0} + a_1\bar{\rho}^{v_1} + a_2\bar{\rho}^{v_2} + \dots$ is positive provided (without loss of generality) $a_0 > 0$. Now, we may write $\bar{\sigma} = \psi\sigma$, where $\sigma = a_0\rho^{v_0} + \tau$, ${}^o(\log_\rho|\tau|) \geq v_1$. It follows that if v is an arbitrary standard real number between v_0 and v_1 , $v_0 < v < v_1$ then $\log_\rho|\tau| > v$, $|\tau| < \rho^v$, $a_0\rho^{v_0} > |\tau|$ and so

$$\sigma = a_0\rho^{v_0} + \tau \geq a_0\rho^{v_0} - |\tau| > 0$$

and hence, $\bar{\sigma} > 0$. Thus Φ is order preserving, as asserted.

5. Functions in ${}^o R$. Let $f(x)$ be any real-valued function defined for $a < x < b$, $a, b \in R$. On passing to *R , $f(x)$ is extended automatically to a function ${}^*f(x)$ which is defined for $a < x < b$ in *R . We wish to find a natural extension of the function $f(x)$ as we pass from R to ${}^o R$.

Such an extension can be obtained, under certain conditions, as follows. Let ξ be any element of ${}^o R$ between a and b , $a < \xi < b$. Let ψ be the canonical homomorphism from M_0 to ${}^o R$ as before (see (4.2) above). Then we define

$$(5.1) \quad {}^o f(\xi) = \psi({}^*f(x)) \text{ for } x \in \psi^{-1}\xi, \quad a < x < b$$

provided the expression on the right hand side of (5.1) is independent of the particular choice of x subject to the stated conditions ($a < x < b$, $\psi x = \xi$).

Suppose in particular that $f(x)$ satisfies a Lipschitz condition in any closed subinterval of $a < x < b$. Thus, for any $a < a' < b' < b$ there exists a $k = k(a', b')$ such that for any $a' \leq x_1 < x_2 \leq b'$,

$$(5.2) \quad |f(x_2) - f(x_1)| \leq k|x_2 - x_1|.$$

Passing from R to *R , we see that (5.2) still holds, for standard a' , b' and for arbitrary x_1 , x_2 in the interval $\langle a', b' \rangle$, if we affix a star to $f(x_2)$ and $f(x_1)$. In particular, it therefore holds for two points x_1 , x_2 of *R which are infinitely close to some standard x_0 , $a < x_0 < b$ (where the constant k may depend on x_0).

Now let $\xi \in {}^o R$ be infinitely close to $x_0 \in R$. Then if x_1 , x_2 belong to $\psi^{-1}\xi$, both x_1 and x_2 are infinitely close to x_0 in *R , and (5.2) applies for an appropriate standard k . But then $x_2 - x_1 \in M_1$ and so, by (5.2), $f(x_2) - f(x_1) \in M_1$, $\psi f(x_2) = \psi f(x_1)$. This shows that in this case, (5.1) provides a unique definition for ${}^o f(\xi)$.

In particular, the Lipschitz condition is satisfied if $f(x)$ has a continuous derivative for $a < x < b$ or, more particularly, if $f(x)$ is infinitely differentiable in that interval. Suppose that this is the case and consider the restriction of ${}^o f(x)$ to points

$$\xi = a_0 + a_1\bar{\rho}^{v_1} + a_2\bar{\rho}^{v_2} + \dots, \quad 0 < v_1 < v_2 < \dots, \quad a < a_0 < b.$$

We may compare ${}^{\rho}f(x)$ for such a point with the function which is obtained by transferring Laugwitz' definition from L to ${}^{\rho}L$, i.e., with the function

$$F(x) = \Phi({}^L f(\Phi^{-1}x)).$$

We propose to show that ${}^{\rho}f(x)$ actually *coincides* with $F(x)$ for such argument values,

$$(5.3) \quad {}^{\rho}f(\xi) = \Phi({}^L f(\Phi^{-1}\xi)).$$

In order to verify this identity, we observe that, except for rearrangements (which can be justified without difficulty within ${}^{\rho}R$), the right hand side of (5.3) is simply the formal Taylor expansion in ${}^{\rho}R$ of $f(x)$ about the point a_0 . Thus, our claim is that

$$(5.4) \quad \begin{aligned} {}^{\rho}f(\xi) = & f(a_0) + \frac{f'(a_0)}{1!}(a_1\bar{\rho}^{v_1} + a_2\bar{\rho}^{v_2} + \dots) + \frac{f''(a_0)}{2!}(a_1\bar{\rho}^{v_1} + a_2\bar{\rho}^{v_2} + \dots)^2 \\ & + \dots + \frac{f^{(n)}(a_0)}{n!}(a_1\bar{\rho}^{v_1} + a_2\bar{\rho}^{v_2} + \dots)^n + \dots, \end{aligned}$$

in other words, that the Taylor series of ${}^{\rho}f$ about a_0 converges at ξ . Put $\eta = a_1\bar{\rho}^{v_1} + a_2\bar{\rho}^{v_2} + \dots$ and choose $h \in \psi^{-1}\eta$, so that $a_0 + h \in \psi^{-1}\xi$. By Taylor's formula with Lagrange's remainder term

$$\begin{aligned} {}^*f(a_0 + h) = & {}^*f(a_0) + \frac{{}^*f'(a_0)}{1!}h + \frac{{}^*f''(a_0)}{2!}h^2 + \dots \\ & + \frac{{}^*f^{(n)}(a_0)}{n!}h^n + \frac{{}^*f^{(n+1)}(a_0 + \theta h)}{(n+1)!}h^{n+1}, \end{aligned}$$

where $0 \leq \theta \leq 1$. Now on the right hand side of this identity ${}^*f^{(k)}(a_0) = f^{(k)}(a_0)$ since a_0 is standard. Also, since $f(x)$ is infinitely differentiable, $f^{(n+1)}(x)$, and hence ${}^*f^{(n+1)}(x)$, is bounded by a standard real number in any standard closed subinterval of $\langle a, b \rangle$ and hence, is bounded by a standard number B in the monad of a_0 . Hence

$$(5.5) \quad \left| \frac{{}^*f^{(n+1)}(a_0 + \theta h)}{(n+1)!} h^{n+1} \right| \leq B |h|^{n+1}.$$

Let v be any standard positive number less than v_1 . Then (5.5) together with the fact that $v(\eta) = {}^o(\log_{\rho}|h|) \geq v_1$ shows that

$$\left| \frac{{}^*f^{(n+1)}(a_0 + \theta h)}{(n+1)!} h^{n+1} \right| < \rho^{(n+1)v}.$$

Hence

$$\left| {}^*f(a_0 + h) - \sum_{k=0}^n \frac{f^{(k)}(a_0)}{k!} h^k \right| < \rho^{(n+1)v},$$

and so

$$v\left({}^{\rho}f(\xi) - \sum_{k=0}^n \frac{f^{(k)}(a_0)}{k!} \eta^k\right) \geq (n+1)v.$$

This shows that

$${}^{\rho}f(\xi) = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{f^{(k)}(a_0)}{k!} \eta^k,$$

proving (5.4).

The identity (5.3) is of interest in itself since it provides a natural justification of Laugwitz' definition within a more comprehensive framework. Beyond that, by relating Laugwitz' theory to that wider framework, we are able to make use of the resources of Nonstandard Analysis in order to provide satisfactory answers to several problems which were left open by Laugwitz. We shall turn to this task in our next section.

6. The intermediate value theorem in L . In view of (5.3), the function ${}^{\rho}f(x)$, with values restricted to ${}^{\rho}L$, behaves in exactly the same way as the function ${}^L f(x)$ on a corresponding interval in L . Consider the real valued function $f(x)$ which is defined in the interval $-1 < x < 1$ by

$$(6.1) \quad f(x) = \begin{cases} e^{-1/|x|} & \text{for } x \neq 0, \\ 0 & \text{for } x = 0. \end{cases}$$

Then $f(x)$ is infinitely differentiable in the entire interval of definition, including $x = 0$. At that point $f^{(n)}(x) = 0$ for $n = 0, 1, 2, \dots$.

Let $x_1 = 0$, $x_2 = \frac{1}{2}$. Then ${}^{\rho}f(x_1) = f(x_1) = 0$, ${}^{\rho}f(x_2) = f(x_2) = 1/e^2$. If ${}^{\rho}f(x)$ satisfied the intermediate value theorem, there would exist a $\xi \in {}^{\rho}L$, $0 < \xi < \frac{1}{2}$, such that ${}^{\rho}f(\xi) = \bar{\rho}$. We shall show that there is no such ξ .

Suppose first that ξ is infinitely close to 0,

$$\xi = a_0 \bar{\rho}^{v_0} + a_1 \bar{\rho}^{v_1} + a_2 \bar{\rho}^{v_2} + \dots, \quad 0 < v_0 < v_1 < \dots \rightarrow \infty, \quad a_0 > 0.$$

Then, by (5.4)

$${}^{\rho}f(\xi) = f(0) + \frac{f'(0)}{1!} \xi + \frac{f''(0)}{2!} \xi^2 + \dots = 0$$

so ${}^{\rho}f(\xi)$ cannot be equal to $\bar{\rho}$.

Suppose next that ξ is not infinitely close to 0. Then $\xi = a_0 + \eta$, where $0 < a_0 \leq \frac{1}{2}$, $v(\eta) > 0$ and so, by (5.4), ${}^{\rho}f(\xi) = f(a_0) + \xi$, where $v(\xi) > 0$. This shows that ${}^{\rho}f(\xi)$ is infinitely close to $f(a_0)$, which is a standard real number different from 0, and so again ${}^{\rho}f(\xi)$ cannot be equal to $\bar{\rho}$, which is infinitesimal.

By contrast, if $f(x)$ is continuous in an interval $a < x < b$ and if the definition

(5.1) is effective in an interval $x_1 \leq x \leq x_2$ where $a < x_1 < x_2 < b$, $x_1, x_2 \in {}^oR$, then the intermediate value theorem does apply in oR . That is to say, under these conditions:

6.2. THEOREM. *If ${}^o f(x_1) < \eta < {}^o f(x_2)$ for $\eta \in {}^oR$, then there exists a $\xi \in {}^oR$, $x_1 < \xi < x_2$ such that ${}^o f(\xi) = \eta$.*

To see this, we only have to choose elements of $*R$, x'_1, x'_2, η' such that $\psi x'_1 = x_1$, $\psi x'_2 = x_2$, $\psi \eta' = \eta$. Then $*f(x'_1) < \eta' < *f(x'_2)$ and so, by the intermediate value theorem for $*f(x)$ there exists a $\xi' \in *R$, $x'_1 < \xi' < x'_2$ such that $*f(\xi') = \eta'$. Putting $\xi = \psi \xi'$ we then have ${}^o f(\xi) = \psi(*f(\xi')) = \psi \eta' = \eta$. This shows that the intermediate value theorem is satisfied in this case.

For the remainder of this section, it will be our main purpose to show that the intermediate value theorem holds also in oL for functions ${}^o f(x)$ which are obtained from infinitely differentiable functions $f(x)$ in R —and hence, holds also in L for the corresponding functions ${}^L f(x)$ —subject to rather mild restrictions, as follows.

6.3. THEOREM. *Let $f(x)$ be a real-valued function which is defined and infinitely differentiable for $a < x < b$, $a, b \in R$ and let ${}^o f(x)$ be defined by 5.1. Suppose that for every $x \in R$, $a < x < b$, there is a positive integer n such that $f^{(n)}(x) \neq 0$. For any $x_1, x_2 \in {}^oL$, $a < x_1 < x_2 < b$, let a_0 and b_0 be the uniquely defined elements of R which are infinitely close to x_1 and x_2 respectively, i.e.,*

$$x_1 = a_0 + a_1 \bar{\rho}^{v_1} + a_2 \bar{\rho}^{v_2} + \dots, \quad 0 < v_1 < v_2 < \dots \rightarrow \infty,$$

$$x_2 = b_0 + b_1 \bar{\rho}^{\mu_1} + b_2 \bar{\rho}^{\mu_2} + \dots, \quad 0 < \mu_1 < \mu_2 < \dots \rightarrow \infty,$$

and suppose that $a < a_0 \leq b_0 < b$. Let η be an element of oL such that ${}^o f(x_1) < \eta < {}^o f(x_2)$.

Then there exists a $\xi \in {}^oL$, $x_1 < \xi < x_2$, such that ${}^o f(\xi) = \eta$.

Proof. Comparing 6.3 with 6.2 (which applies to the situation described in 6.3) we see that we only have to show that the $\xi \in {}^oR$ mentioned in 6.2 belongs more particularly to oL . Choosing x'_1, x'_2, η' as in the proof of 6.2 such that $\psi x'_1 = x_1$, $\psi x'_2 = x_2$, $\psi \eta' = \eta$ we have, for some $\xi' \in *R$, $x'_1 < \xi' < x'_2$, $*f(\xi') = \eta'$ and hence ${}^o f(\xi) = \eta$ where $\xi = \psi \xi'$. Now $x'_1 < \xi' < x'_2$ implies that ξ' is finite and has a standard part, ${}^o \xi' = d_0$, where $a < a_0 \leq d_0 \leq b_0 < b$. At the same time, η must be of the form $e_0 + e_1 \bar{\rho}^{\lambda_1} + e_2 \bar{\rho}^{\lambda_2} + \dots$, $0 < \lambda_1 < \lambda_2 < \dots \rightarrow \infty$ since it is in oL and finite. Hence, ${}^o \eta' = e_0$ and $f(d_0) = e_0$.

Suppose now that $f'(d_0) \neq 0$. Then the inversion theorem is applicable. It follows that there exist $h_1 > 0$, $h_2 > 0$, $k_1 > 0$, $k_2 > 0$, such that $f(x)$ is a one-to-one mapping of the interval D defined by $d_0 - h_1 < x < d_0 + h_2$ on the interval E defined by $e_0 - k_1 < y < e_0 + k_2$ such that the inverse function $g(y) = f^{-1}(y)$ is infinitely differentiable on E . Passing to $*R$, we find that the infinitely differentiable function $*f(x)$ maps $*D$ in one-to-one correspondence on $*E$ such that $*g(y)$ is the inverse of

this mapping and is infinitely differentiable as well (in the sense of $*R$). Hence, $*f(\xi') = \eta'$ entails $*g(\eta') = \xi'$ and so

$$\xi = \psi\xi' = \psi(*g(\eta')) = {}^p g(\eta) \in {}^p L,$$

proving our assertion in this case.

Dropping the restriction that $f'(d_0) \neq 0$ (but not excluding this case) we put $F(x) = f(x) - f(d_0)$ and define $H(x)$ for $a < x < b$ by

$$H(x) = \begin{cases} \frac{F(x)}{x - d_0} & \text{for } x \neq d_0 \\ F'(d_0) & \text{for } x = d_0. \end{cases}$$

Also, on the assumption of our theorem, there is an $n \geq 0$ such that

$$F(d_0) = F'(d_0) = \dots = F^{(n)}(d_0) = 0, \quad F^{(n+1)}(d_0) \neq 0.$$

Then $F(x) = H(x)(x - d_0)$, and so

$$(6.4) \quad F'(x) = H(x) + H'(x)(x - d_0) \quad \text{for } x \neq d_0$$

and, more generally,

$$(6.5) \quad F^{(k)}(x) = kH^{(k-1)}(x) + H^{(k)}(x)(x - d_0)$$

for $k = 1, 2, \dots$, $x \neq d_0$, $a < x < b$, by induction.

We now wish to show that, for $x \neq d_0$,

$$(6.6) \quad H^{(\lambda)}(x) = \frac{F^{(\lambda+1)}(d_0)}{\lambda + 1} + \frac{F^{(\lambda+2)}(d_0)}{1!(\lambda + 2)}(x - d_0) + \dots \\ + \frac{F^{(\lambda+m)}(d_0)}{(m-1)!(\lambda + m)}(x - d_0)^{m-1} + G_{\lambda,m}(x - d_0)^m$$

provided $\lambda \geq 1$, $m \geq 1$, where $G_{\lambda,m}$ is a linear combination with rational coefficient of values of $F^{(\lambda+m+1)}(x)$ taken at points x' in the interval $\langle d_0, x \rangle$.

For $\lambda = 1$, we have the Taylor expansion for $F'(x)$

$$F'(x) = F'(d_0) + \frac{F''(d_0)}{1!}(x - d_0) + \dots + \frac{F^{(1+m)}(d_0)}{m!}(x - d_0)^m \\ + \frac{F^{(2+m)}(d_0 + \theta_1(x - d_0))}{(m+1)!}(x - d_0)^{m+1}$$

where $0 \leq \theta_1 \leq 1$, while the Taylor expansion for $F(x)$ yields

$$(6.7) \quad H(x) = F'(d_0) + \frac{F''(d_0)}{2!}(x - d_0) + \dots + \frac{F^{(1+m)}(d_0)}{(m+1)!}(x - d_0)^m \\ + \frac{F^{(2+m)}(d_0 + \theta_0(x - d_0))}{(m+2)!}(x - d_0)^{m+1},$$

where $0 \leq \theta_0 \leq 1$. Hence, from (6.4),

$$\begin{aligned} H'(x) &= \frac{F'(x) - H(x)}{(x - d_0)} \\ &= \frac{F''(d_0)}{2} + \cdots + \frac{F^{(1+m)}(d_0)}{(m-1)!(1+m)}(x - d_0)^{m-1} + G_{1,m}(x - d_0)^m, \end{aligned}$$

where

$$G_{1,m} = \frac{F^{(2+m)}(d_0 + \theta_1(x - d_0))}{(m+1)!} - \frac{F^{(2+m)}(d_0 + \theta_0(x - d_0))}{(m+2)!},$$

as required.

Suppose now that the assertion has been proved up to some $\lambda \geq 1$, for all $m \geq 1$. In order to prove the corresponding formula for $\lambda + 1$, we write down the appropriate Taylor expansion for $F^{(\lambda+1)}(x)$, so

$$\begin{aligned} F^{(\lambda+1)}(x) &= F^{(\lambda+1)}(d_0) + \frac{F^{(\lambda+2)}(d_0)}{1!}(x - d_0) + \cdots + \frac{F^{(\lambda+m+1)}(d_0)}{m!}(x - d_0)^m \\ &\quad + \frac{F^{(\lambda+m+2)}(d_0 + \theta_{\lambda+1}(x - d_0))}{(m+1)!}(x - d_0)^{m+1}, \end{aligned}$$

where $0 \leq \theta_{\lambda+1} \leq 1$. Then, by (6.5) and (6.6) (with $m+1$ for m)

$$\begin{aligned} H^{(\lambda+1)}(x) &= \frac{F^{(\lambda+1)}(x) - (\lambda+1)H^{(\lambda)}(x)}{x - d_0} \\ &= \frac{F^{(\lambda+2)}(d_0)}{\lambda+2} + \cdots + \frac{F^{(\lambda+m+1)}(d_0)}{(m-1)!(\lambda+1+m)}(x - d_0)^{m-1} + G_{\lambda+1,m}(x - d_0)^m, \end{aligned}$$

where

$$G_{\lambda+1,m} = \frac{F^{(\lambda+m+2)}(d_0 + \theta_{\lambda+1}(x - d_0))}{(m+1)!} - (\lambda+1)G_{\lambda,m+1}.$$

This proves (6.6). We now obtain immediately, for $\lambda \geq 1$

$$(6.8) \quad \lim_{x \rightarrow d_0} H^{(\lambda)}(x) = \frac{F^{(\lambda+1)}(d_0)}{\lambda+1}$$

and this is true also for $\lambda = 0$, by (6.7). On the other hand, we may calculate the derivatives of $H(x)$ at d_0 . We have, from (6.7), which is valid also for $m = 0$,

$$\begin{aligned} H'(d_0) &= \lim_{x \rightarrow d_0} \frac{H(x) - H(d_0)}{x - d_0} = \lim_{x \rightarrow d_0} \frac{H(x) - F'(d_0)}{x - d_0} = \lim_{x \rightarrow d_0} \frac{F''(d_0 + \theta_0(x - d_0))}{2} \\ &= \frac{F''(d_0)}{2}, \end{aligned}$$

where θ_0 may depend on x . Thus, $H(x)$ has a continuous derivative everywhere in its interval of definition.

Suppose now that we have proved that $H(x)$ has continuous derivatives up to order $\lambda \geq 1$ in the entire interval of definition $a < x < b$ such that $H^{(\lambda)}(d_0) = F^{(\lambda+1)}(d_0)/(\lambda+1)$. We then make use of (6.6) for $m=2$, where we observe that $G_{\lambda,2}$ (as a linear combination with fixed rational coefficients of values of $F^{(\lambda+3)}$ for arguments x' in the interval $\langle d_0, x \rangle$) remains bounded in the neighborhood of x_0 . Hence, for such x ,

$$H^{(\lambda)}(x) = \frac{F^{(\lambda+1)}(d_0)}{\lambda+1} + \frac{F^{(\lambda+2)}(d_0)}{\lambda+2}(x-d_0) + O(x-d_0)^2$$

and so

$$\lim_{x \rightarrow d_0} \frac{H^{(\lambda)}(x) - H^{(\lambda)}(d_0)}{x - d_0} = \frac{F^{(\lambda+2)}(d_0)}{\lambda+2} = \lim_{x \rightarrow d_0} H^{(\lambda+1)}(x).$$

This shows that $H(x)$ possesses continuous derivatives of all orders in its interval of definition. In particular

$$(6.9) \quad H^{(\lambda)}(d_0) = F^{(\lambda+1)}(d_0)/(\lambda+1), \quad \lambda = 0, 1, \dots$$

and so

$$H(d_0) = H'(d_0) = \dots = H^{(n-1)}(d_0) = 0, \quad H^{(n)}(d_0) = \frac{F^{(n+1)}(d_0)}{n+1} \neq 0.$$

If $n > 0$, we may repeat our procedure, obtaining from $H(x)$ a function $H_1(x)$ in the same way in which we obtained $H(x)$ from $F(x)$. Thus, putting

$$H_1(x) = \begin{cases} H(x)/(x-d_0) = F(x)/(x-d_0)^2 & \text{for } x \neq d_0 \\ H'(d_0) & \text{for } x = d_0, \end{cases}$$

we find that $H_1(x)$ is infinitely differentiable for $a < x < b$ and

$$H_1(d_0) = H'_1(d_0) = \dots = H_1^{(n-2)}(d_0) = 0, \quad H_1^{(n-1)}(d_0) = \frac{F^{(n+1)}(d_0)}{n(n+1)} \neq 0.$$

Continuing in this way, we obtain after $n-1$ more steps the function

$$H_n(x) = \begin{cases} H_{n-1}(x)/(x-d_0) = F(x)/(x-d_0)^{n+1} = \frac{f(x)-f(d_0)}{(x-d_0)^{n+1}}, & \text{for } x \neq d_0 \\ \frac{f^{(n+1)}(d_0)}{(n+1)!} \neq 0, & \text{for } x = d_0 \end{cases}$$

which is infinitely differentiable for $a < x < b$.

Suppose first that n is even, $n+1$ is odd. Then the function $w^{1/(n+1)}$, with

the determination that $(H_n(d_0))^{1/(n+1)}$ be real, is infinitely differentiable in the neighborhood of $H_n(d_0)$ and so the function $P(x) = (H_n(x))^{1/(n+1)}$ is infinitely differentiable in some neighborhood of $x = d_0$, for $d_0 - h < x < d_0 + h$, say. The function

$$Q(x) = P(x)(x - d_0) = (f(x) - f(d_0))^{1/(n+1)}$$

therefore is also infinitely differentiable in the same interval, and

$$Q'(x) = P(x) + P'(x)(x - d_0), \quad Q'(d_0) = P(d_0) \neq 0.$$

Passing to *R we see that, for $x = \xi'$,

$${}^*Q(\xi') = (f(\xi') - f(d_0))^{1/(n+1)} = (\eta' - e_0)^{1/(n+1)}.$$

Hence

$${}^{\rho}Q(\xi) = \psi((\eta' - e_0)^{1/(n+1)}) = (\eta - e_0)^{1/(n+1)} \in {}^{\rho}L$$

since L , and hence ${}^{\rho}L$, is real closed (see section 3 above). Hence, applying the inversion theorem to $Q(x)$ at $x = d_0$ (exactly as we applied it earlier to $f(x)$ on the assumption that $f'(d_0) \neq 0$), and letting $S(y)$ be the inverse function to $Q(x)$ at $x = d_0$, $y = 0$, we obtain

$$\xi = \psi \xi' = \psi({}^*S((\eta' - e_0)^{1/(n+1)})) = {}^{\rho}S(\eta - e_0)^{1/(n+1)} \in {}^{\rho}L.$$

This disposes of the case that n is even.

Suppose finally that n is odd, $n + 1$ is even. We may assume without loss of generality that $H_n(d_0) = f^{(n+1)}(d_0)/(n+1)!$ is positive, otherwise we consider $-f(x)$ in place of $f(x)$. Then $f(x) - f(d_0) = H_n(x)(x - d_0)^{n+1}$ must be positive, for $x \neq d_0$ in a sufficiently small neighborhood of d_0 . Introducing $P(x) = (H_n(x))^{1/(n+1)}$ with the positive determination for $(H_n(d_0))^{1/(n+1)}$, and $Q(x) = P(x)(x - d_0)$ we then have again that $P(x)$ and $Q(x)$ are infinitely differentiable in a neighborhood of $x = d_0$, and that $Q'(d_0) = P(d_0) \neq 0$. Also,

$${}^*Q(\xi') = \pm (f(\xi') - f(d_0))^{1/(n+1)} = \pm (\eta' - e_0)^{1/(n+1)}$$

leading to ${}^{\rho}Q(\xi) = \pm (\eta - e_0)^{1/(n+1)}$, which is again an element of ${}^{\rho}L$. Finally, introducing the inverse function $S(y)$ of $Q(x)$ with $S(0) = d_0$, as before, we have

$$\xi = \psi \xi' = \psi(S(\pm (\eta' - e_0)^{1/(n+1)})) = {}^{\rho}S(\pm (\eta - e_0)^{1/(n+1)}) \in {}^{\rho}L.$$

The proof of Theorem 6.3 is now complete.

Although the counterexample given at the beginning of the section shows that some restriction on the behavior of the derivatives of $f(x)$ is required, the particular set of conditions given in 6.3, is not strictly necessary. Thus, if $f(x) = \text{const.}$, then the conditions of the theorem are not satisfied but its conclusion is, trivially. Nevertheless, 6.3 includes a large number of interesting cases, e.g., all non-constant real analytic functions $f(x)$.

7. The mean value theorem. Suppose the function $f(x)$ is continuously differentiable for $a < x < b$. Let D be the set of points $\xi \in {}^pR$ such that ξ is infinitely close to a point a_0 in the interior of that interval, $a < a_0 < b$. Then $f'(x)$ is bounded in any closed subinterval $a' \leq x \leq b'$ of $a < x < b$ and so $f(x)$ satisfies a Lipschitz condition in that interval. Taking $a' < a_0 < b'$ we see, therefore, that the definition 5.1 is effective. We claim, moreover, that the resulting function ${}^pf(x)$ is continuous, in the sense of the metric of pR , at all points $\xi \in D$.

To see this, let $\{\xi_n\}$ be a sequence of elements of D such that $\lim_{n \rightarrow \infty} \xi_n = \xi$ and choose a number ξ' and a sequence $\{\xi'_n\}$ in *R such that $\psi\xi' = \xi$, $\psi\xi'_n = \xi_n$, $n=0, 1, 2, \dots$. Since $\lim \xi_n = \xi$, there exist $a', b' \in R$ such that $a' < \xi < b'$, $a' < \xi' < b'$, $a' < \xi_n < b'$, $a' < \xi'_n < b'$, $n=0, 1, 2, \dots$. Let m be a bound for $f'(x)$ in the closed interval $a' \leq x \leq b'$ within R and, hence within *R . Then

$${}^*f(\xi'_n) - {}^*f(\xi') = {}^*f'(\xi' + \theta(\xi'_n - \xi'))(\xi'_n - \xi')$$

for some $0 \leq \theta \leq 1$ and, hence

$$|{}^pf(\xi_n) - {}^pf(\xi)| \leq m |\xi_n - \xi|.$$

This, together with $\lim \xi_n = \xi$ implies $\lim {}^pf(\xi_n) = {}^pf(\xi)$, proving our assertion.

Suppose next that $f(x)$ is twice continuously differentiable for $a < x < b$. In this case, we propose to show that ${}^pf(x)$ is differentiable in D (in the sense of the metric of pR) and that on D ,

$$(7.1) \quad \frac{d}{dx} {}^pf(x) = {}^p(f'(x)).$$

For ξ in D and $\eta \neq 0$ such that $\xi + \eta$ also belongs to D , choose ξ' and η' for which $\psi\xi' = \xi$, $\psi\eta' = \eta$. Then there exists a $\theta' \in {}^*R$, $0 \leq \theta' \leq 1$, such that

$$(7.2) \quad \frac{f(\xi' + \eta') - f(\xi')}{\eta'} = f'(\xi' + \theta'\eta').$$

Applying the mapping ψ to (7.2), we obtain

$$(7.3) \quad \frac{{}^pf(\xi + \eta) - {}^pf(\xi)}{\eta} = ({}^pf'(x))_{x=\xi+\theta\eta},$$

where $\theta = \psi\theta'$. Now let η tend to zero. Then the right hand side of (7.3) tends to $({}^pf'(x))_{x=\xi}$ since ${}^pf'(x)$ is continuous on D . This proves (7.1).

In particular, if $f(x)$ is infinitely differentiable, then ${}^pf(\xi)$ and $({}^pf'(x))_{x=\xi}$ belong to pL for $\xi \in {}^pL$. It follows that in that case ${}^pf(x)$ is defined and infinitely differentiable in $D \cap {}^pL$. Accordingly, the same is true of the function

$${}^L f(x) \text{ for } x = a_0 + a_1 t^{v_1} + a_2 t^{v_2} + \dots, \quad 0 < v_1 < v_2 < \dots \rightarrow \infty, \quad a < a_0 < b.$$

(7.3), in combination with (7.1) shows also that the mean value theorem holds in

pR under the stated conditions, more particularly for infinitely differentiable $f(x)$. However, here again we may show that the mean value theorem breaks down, for certain infinitely differentiable functions, both in pL and in L . The function (6.1) which provided an example for the breakdown of the intermediate value theorem, will do also for the present issue as can be seen by considering the ratio of increments $(f(\xi_2) - f(\xi_1))/(\xi_2 - \xi_1)$ for $\xi_2 = 1/(2 + \bar{\rho})$, $\xi_1 = -\frac{1}{2}$. There is no $\xi_3 \in {}^pL$ in the closed interval from ξ_1 to ξ_2 such that $({}^pf(x))'$ is equal to that ratio for $x = \xi_3$.

We shall prove, as our principal positive result in this area:

7.4. THEOREM. *Let $f(x)$ be a real valued function which is defined and infinitely differentiable for $a < x < b$; $a, b \in R$ and let ${}^pf(x)$ be defined by 5.1. Suppose that for every x , $a < x < b$, there is an integer $n \geq 2$ such that $f^{(n)}(x) \neq 0$. For any $x_1, x_2 \in {}^pL$, $a < x_1 < x_2 < b$, let a_0 and b_0 be the uniquely defined elements of R which are infinitely close to x_1 and x_2 respectively, i.e.,*

$$x_1 = a_0 + a_1\bar{\rho}^{v_1} + a_2\bar{\rho}^{v_2} + \cdots, \quad 0 < v_1 < v_2 < \cdots \rightarrow \infty,$$

$$x_2 = b_0 + b_1\bar{\rho}^{\mu_1} + b_2\bar{\rho}^{\mu_2} + \cdots, \quad 0 < \mu_1 < \mu_2 < \cdots \rightarrow \infty,$$

and suppose that $a < a_0 \leq b_0 < b$.

Then there exists a $\xi \in {}^pL$, $x_1 \leq \xi \leq x_2$ such that

$$\frac{{}^pf(x_2) - {}^pf(x_1)}{x_2 - x_1} = \left(\frac{d}{dx} {}^pf(x) \right)_{x=\xi}.$$

Here again there is an exactly corresponding theorem for the function ${}^Lf(x)$ in L . The conditions of the theorem are not necessary since they exclude all functions of constant slope, for which the conclusion of the theorem is obviously correct. However, the theorem is nevertheless of a rather general character, including, for example, all other real analytic functions.

For the proof, we require the following auxiliary consideration.

Assume that the conditions of (7.4) are satisfied and choose $x'_1 \in \psi^{-1}x_1$, $x'_2 \in \psi^{-1}x_2$. Then we claim that $*f'(x)$ attains its maximum in the interval $x'_1 \leq x \leq x'_2$ either at x'_1 or at x'_2 or at some standard point x_0 , $x'_1 \leq x_0 \leq x'_2$ (but, possibly, also elsewhere).

Suppose that $*f(x)$ attains its maximum neither at x'_1 nor at x'_2 but at a point \bar{x} , $x'_1 < \bar{x} < x'_2$. Let x_0 be the standard part of \bar{x} , $x_0 = {}^0\bar{x}$. Suppose that $x_0 < x_1$ (so that $x_0 = a_0$). Depending on whether the first non-vanishing derivative of $f'(x)$ at x_0 is either positive or negative, $*f'(x)$ will be either strictly increasing or strictly decreasing in some interval $x_0 \leq x \leq x_0 + h$, where h is standard and positive. Since \bar{x} and x'_1 belong to that interval, the latter case would involve $*f'(x'_1) > *f'(\bar{x})$, contrary to our choice of \bar{x} . Accordingly, we have to assume that $*f'(x)$ increases strictly for $x_0 \leq x \leq x_0 + h$. Now x'_2 cannot belong to that interval for then $*f'(x'_2) > *f'(\bar{x})$, which is again impossible. It follows that $\bar{x} < x_0 + h < x'_2$ and

$*f'(x_0 + h) > f(\bar{x})$ which is also impossible. We therefore conclude that $x_0 \geq x_1$ and, by similar reasoning, $x_0 \leq x_2$. The discussion of the variation of $*f'(x)$ in the neighborhood of x_0 shows that we must exclude both $x_0 < \bar{x}$ and $x_0 > \bar{x}$ and so we conclude that $\bar{x} = x_0$.

Thus we have shown that $*f'(x)$ attains its maximum at x'_1 or at x'_2 or at some standard point $x'_1 < x_0 < x'_2$ (although several of these cases may occur simultaneously). Accordingly $*f'(x)$ attains its maximum in the interval $x'_1 \leq x \leq x'_2$ in all cases at a point ζ'_2 such that $\psi\zeta'_2 = \zeta_2 \in {}^pL$. By a similar argument (or, by applying the conclusion to $-f(x)$) we find that $*f'(x)$ attains its minimum in the same interval at a point ζ'_1 such that $\psi\zeta'_1 = \zeta_1 \in {}^pL$. Passing from $*R$ to pR , we then conclude that ${}^p(f'(x))$ attains its maximum and minimum in $x_1 \leq x \leq x_2$ at points ζ_1 and ζ_2 which belong to pL .

By a well-known formula of the Integral Calculus, which can be transferred from R to $*R$, we have

$$*f'(\zeta'_2)(x'_2 - x'_1) \leq \int_{x'_1}^{x'_2} *f'(x)dx \leq *f'(\zeta'_1)(x'_2 - x'_1),$$

i.e.,

$$*f'(\zeta'_2)(x'_2 - x'_1) \leq *f(x'_2) - *f(x'_1) \leq *f'(\zeta'_1)(x'_2 - x'_1).$$

We apply the mapping ψ to this chain of inequalities and obtain

$${}^p(f'(\zeta_2))(x_2 - x_1) \leq {}^pf(x_2) - {}^pf(x_1) \leq {}^p(f'(\zeta_1))(x_2 - x_1)$$

or, equivalently

$${}^p(f'(\zeta_2)) \leq \frac{{}^pf(x_2) - {}^pf(x_1)}{x_2 - x_1} \leq {}^p(f'(\zeta_1)).$$

But this shows that $({}^pf(x_2) - {}^pf(x_1))/(x_2 - x_1)$ is intermediate between ${}^p(f'(\zeta_2))$ and ${}^p(f'(\zeta_1))$. Hence, by the intermediate value Theorem 6.3, there exists a $\xi \in {}^pL$ which belongs to the closed interval with endpoint ζ_1 and ζ_2 and, hence, belongs to $x_1 \leq x \leq x_2$, such that

$$\frac{{}^pf(x_2) - {}^pf(x_1)}{x_2 - x_1} = {}^p(f'(\xi))$$

and this is the same as

$$\frac{{}^pf(x_2) - {}^pf(x_1)}{x_2 - x_1} = \left(\frac{d}{dx} {}^pf(x) \right)_{x=\xi},$$

by (7.1). The proof of 7.4 is now complete.

8. Conclusion. As Laugwitz points out, his method for extending a function $f(x)$ from R to L applies only in the infinitesimal neighborhood of a point at which $f(x)$ is infinitely differentiable and hence, possesses at least a formal Taylor series. However, if we consider points in the infinitesimal neighborhood of the endpoints of the interval of definition $a < x < b$ of $f(x)$, e.g., $x = a + a_1 t^{v_1} + a_2 t^{v_2} + \dots$, $0 < v_1 < v_2 < \dots$, $a_1 > 0$, then we can still define ${}^L f$ at x , provided f possesses an asymptotic expansion at $x = a$ as x tends to a from the right. Similarly, if $f(x)$ is defined in a semi-infinite interval, for $x > a$ say, we can define ${}^L f(x)$ for positive infinite x provided $f(x)$ possesses an asymptotic expansion as $x \rightarrow +\infty$. In all of these cases, ${}^L f(x)$ can again be obtained "automatically" as $\Phi^{-1}({}^p f(\Phi x))$ (see (5.3) above). However, ${}^p f(x)$ exists also in many cases where no asymptotic expansion as a generalized power series is available, e.g., ${}^p \log x$ exists for positive infinitesimal and infinite x . Conversely, we may investigate the asymptotic expansion of a function $f(x)$ at a singular point (even when it contains logarithmic terms, as happens frequently in the theory of ordinary differential equations) by means of the function ${}^p f(x)$. Going further in the direction of concrete applications, ${}^p R$ also provides us with a convenient framework for the discussion of matched asymptotic expansions for the solution of singular perturbation problems.

This research was supported in part by the National Science Foundation Grant No. GP-18728.

References

1. N. Jacobson, *Lectures in Abstract Algebra*, vol. III, Princeton-Toronto-New-York-London, 1964.
2. D. Laugwitz, Eine nichtarchimedische Erweiterung angeordneter Körper, *Math. Nachr.*, 37 (1968) 225–236.
3. T. Levi-Civita, Sugli infiniti ed infinitesimi attuali quali elementi analitici (1892–1893), *Opere matematiche*, vol. 1, Bologna 1954, pp. 1–39.
4. W. A. J. Luxemburg, What is Nonstandard Analysis, California Institute of Technology, 1968, to be published.
5. A. Ostrowski, Untersuchungen zur arithmetischen Theorie der Körper, *Math. Z.*, 39 (1935) 269–404.
6. A. Robinson, *Non-standard Analysis*, *Studies in Logic and the Foundations of Mathematics*, Amsterdam, 1966.
7. B. L. v. der Waerden, *Algebra*, 5th edition, Berlin-Heidelberg-New York, 1966/1967.
8. O. Zariski and P. Samuel, *Commutative Algebra*, vol. 2, Princeton-Toronto-New-York-London, 1960.

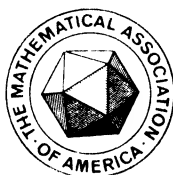
Yale University,
September 1970.

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF
THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 80



NUMBER 7

CODEN: AMMYAE

CONTENTS

Computer Algebra of Polynomials and Rational Functions . . . G. E. COLLINS	725
On the Discrete Version of Wirtinger's Inequality O. SHISHA	755
Current Trends in Algebra GARRETT BIRKHOFF	760

MATHEMATICAL NOTES

Existence of Four Concurrent Normals to a Smooth Closed Hypersurface of E^n BERND WEGNER	782
On a Problem of Besicovitch B. FISHER	785
Topological Properties of the Row Echelon Form G. P. BARKER	787
Two-dimensional Complete Monotonicity with Diagonalization C. H. KIMBERLING	789

RESEARCH PROBLEMS

The Permanent of a Doubly Stochastic Matrix RUSSELL MERRIS	791
----------------------------------------------------------------------	-----

CLASSROOM NOTES

A Generalization of a Theorem of Archimedes WALTER RUDIN	794
The Chromatic Polynomial of a Complete Bipartite Graph . . . J. R. SWENSON	797

MATHEMATICAL EDUCATION

Training Secondary Mathematics Teachers in Venezuela . . . D. B. AICHELE	798
Experiences with Lectures on the History of Mathematics in Utrecht A. F. MONNA	803
Developing Countries: A Rejoinder M. A. B. DEAKIN	806

ELEMENTARY PROBLEMS AND SOLUTIONS	807
ADVANCED PROBLEMS AND SOLUTIONS	814

(Continued on inside cover)

AUGUST-SEPTEMBER

1973

REVIEWS	821
NEWS AND NOTICES	839
MATHEMATICAL ASSOCIATION OF AMERICA	841
MAA Publishes Guidelines for Evaluating College Mathematics Programs	841
Disability Income Plan Added to the MAA Group Insurance Program	844
November Meeting of the Indiana Section	844
February Meeting of the Louisiana-Mississippi Section	845
February Meeting of the Northern California Section	846
March Meeting of the Florida Section	847
New Sectional Governors of the Association	848
Announcement of Lester R. Ford Awards	849
The 1973 William Lowell Putnam Mathematical Competition	849
Films Produced by the MAA	849
Calendars of Future Meetings	852

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss.

Backlog: Main Articles 12 months, Math. Notes 13 months, Research Problems 7 months, Classroom Notes 11 months, Math. Education 10 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to ALEX ROSENBERG, Department of Mathematics, Cornell University, Ithaca, N.Y. 14850; NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D.C. 20036.

HARLEY FLANDERS, *Editor*
ALEX ROSENBERG, *Editor-Elect*

ASSOCIATE EDITORS

JOSHUA BARLAZ	J. G. HARVEY	SEYMOUR SCHUSTER
E. R. BERLEKAMP	ERIC S. LANGFORD	J. ARTHUR SEEBACH, Jr.
JANE W. DI PAOLA	P. D. LAX	E. P. STARKE
ROBERT GILMER	ARTHUR MATTUCK	LYNN A. STEEN
RICHARD GUY	M. W. POWNALL	JAMES WENDEL
RAOUL HAILPERN	GIAN-CARLO ROTA	

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are \$12.50. For nonmembers the subscription price is \$18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June-July, August-September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices.

Copyright © The Mathematical Association of America (Incorporated), 1973

PRINTED IN THE UNITED STATES OF AMERICA

COMPUTER ALGEBRA OF POLYNOMIALS AND RATIONAL FUNCTIONS

G. E. COLLINS, University of Wisconsin

1. Introduction. Computer programs are now available for performing many important algebraic processes of interest and utility to pure and applied mathematicians. Also, many interesting mathematical problems arise in the development and analysis of algorithms for use in such programs. As a mathematician-turned-computer-scientist who has contributed to realizing the current state of computer algebra capabilities and who is intrigued with the mathematical problems whose solutions will contribute to further progress, I hope in this brief survey article to impart some of my knowledge about this subject and to transmit some of my enthusiasm for it to other mathematicians.

The kind of computer algebra I will discuss is concerned mainly with polynomials and rational functions, for the following reasons. First, the allotted space does not permit me to be more ambitious. Second, this is the only part of the subject in which I can really be authoritative. Third, this is the part of the subject about which the most is known and on which other parts most depend.

I shall be concerned with polynomials and rational functions in several variables, primarily with rational integer coefficients, but this will also lead to consideration of rational number coefficients and finite field coefficients; algebraic number coefficients will receive some mention as an advanced topic of current research.

Operations on polynomials and rational functions which will be considered include the "arithmetic" operations of addition, subtraction, multiplication and division. I shall show in Section 3, perhaps to the reader's surprise, that even polynomial arithmetic is non-trivial and interesting when the objective is to design and analyze optimal algorithms.

Algorithms for the arithmetic operations on rational functions require an efficient algorithm for polynomial g.c.d. (greatest common divisor) calculation. Research over the last seven years has revealed that there are numerous versions of the "Euclidean algorithm" for polynomials which differ dramatically in their efficiency. Also, within the last four years, "modular" polynomial g.c.d. algorithms have been devised, which depend on use of the Chinese remainder theorem, and which are orders of magnitudes faster than any of the non-modular Euclidean algorithms.

George Collins earned his Ph.D. at Cornell University, under J. Barkley Rosser. He has served at the IBM Scientific Computing Center, New York, IBM Project Vanguard, New York, and Washington D. C., Mathematics and Applications Department, White Plains, New York, IBM Research Center, Yorktown Heights, Cal. Tech., and the University of Wisconsin, where he has recently been the chairman of the Computer Sciences Department. He has just completed a leave of absence as a visiting professor at Stanford University. He has published extensively in Computer Science and is preparing a two volume book called *Algebraic Algorithms*. Editor.

Furthermore, an intimate relationship has been established between polynomial g.c.d. calculation and polynomial resultant calculation. Hence all of Section 4 is devoted to this remarkable success story.

Section 5 is devoted to algorithms for polynomial factorization. This is a more difficult problem than polynomial g.c.d. calculation, but the progress of the last five years has been equally remarkable. Here again “modular” algorithms have been developed, which reduce the problem of factoring a polynomial with integer coefficients to one of factoring a polynomial with finite field coefficients. This time however, Hensel’s p -adic lemma takes the place of the Chinese remainder theorem, and there is no classical counterpart of the Euclidean algorithm. Instead, reliance must be placed on some important new algorithms due to Berlekamp for factoring polynomials in one variable over a finite field. Polynomial g.c.d. calculation also plays an important role.

Section 6 covers a variety of subjects worthy of mention including polynomials with Gaussian integer coefficients, operations on rational functions, including integration, linear algebra over polynomials and rational functions, computing zeros of polynomials using exact arithmetic and algebra, and calculations with algebraic numbers.

Throughout, we make provisions for polynomials with arbitrarily large coefficients. This is natural since computers are now fast enough that any restrictions imposed by the word length of the computer are artificial and unnecessary. It is also important since only the most trivial algebraic operations on polynomials can be performed without generating integer coefficients which are 100 or more decimal digits in length. Frequently the final result will have coefficients of modest size, but obtaining this result requires the generation of polynomials with very much larger coefficients. Thus, Section 2 is devoted to algorithms and computer techniques for arithmetic operations and g.c.d. calculations on “infinite-precision” integers.

Section 2 also provides an introduction, for the non-computer scientist, to the subject of the analysis of algorithms, especially the analysis of the computing time of an algorithm at a level which is independent of any particular computer that might be used to perform the algorithm. There is a growing conviction among many computer scientists that the analysis of algorithms, or the study of “computational complexity,” is the most important and fundamental part of computer science—if not the only part. Although I share in the enthusiasm of this viewpoint, I believe that it overlooks the equally important role of the computer scientist in discovering, designing and synthesizing new algorithms.

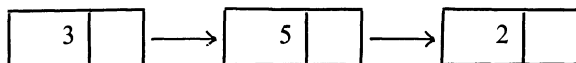
I hope that readers of this article will find it interesting or useful in at least one of the following ways. First, it may create an awareness of the computer programs and systems for computer algebra which are now available and which might be useful as tools in conducting some mathematical research of either a pure or applied nature. Second, the reader may be interested in the mathematical foundations of

the algorithms which are discussed. Third, the reader may be interested in the mathematical methods and problems which arise in attempting to analyze the computing times of these algorithms.

2. Infinite-precision integer arithmetic. One cannot go very far in performing operations on polynomials with rational integer coefficients (hereafter called integral polynomials) without quickly generating polynomials with very large coefficients. Most current computers can directly perform arithmetic on integers no larger than about 10 decimal digits, whereas integers up to several hundred decimal digits are common in various algebraic calculations. The obvious solution is to teach the computer (via subprograms—i.e., auxiliary programs used by other programs) to perform base β arithmetic by the well-known “classical” methods, where β is a “parameter” chosen to suit a particular computer. Thus β will typically be about 10^{10} , but if the computer is itself constructed to perform binary (i.e., base 2) arithmetic, as most are, then there will usually be some advantage in choosing a power of 2 for β . Each “digit” in the base β representation of a large integer is then stored in a separate memory location or “word” of the computer.

The only real problem which arises in this scheme is one of computer memory allocation or, as we also say, storage allocation. Such memory allocation must be done “dynamically,” i.e., when the computer program is executed rather than when the program is written. This is because a program variable may take on a sequence of integers of very different lengths during the course of running the program just once, and it would be very wasteful to allow enough memory for the largest integer which might occur, even if this were possible to predict. The problem is further compounded if we are working with a polynomial some of whose coefficients may be much larger than others (and many of the coefficients may even be zero).

Several satisfactory solutions have been devised for the dynamic storage allocation problem, but the first discovered, the most elegant, and the most universally applicable solution is that of *lists*. Suppose, for illustrative purposes, that $\beta = 10$ and that we wish to store the number 352. We can store the three digits of this number in any three available words of the computer memory, provided we also store in the first word the location, or *address*, of the second word, in the second the address of the third, and in the third an indication, say some standard fictitious address, that there are no further words in this list of words. Since the choice of the three words is insignificant, we represent this diagrammatically as follows:



We say that these three words comprise a representation of the list (3, 5, 2) and that the location of this list is the location of the first word. Each word in the list contains two fields: an *element field* (left half) and a *successor field* (right half). We say that the list (3, 5, 2) *represents* the integer 352.

At any given time, all of the words in some designated portion of the computer memory which are not in use as part of some such data list are themselves all linked together, in arbitrary order, as an *available space list*. When a new word is needed to construct some data list, the first word is unlinked from the available space list and linked to the data list. When a data list is no longer needed to complete a computation, its words or *cells* are linked to the head of the available space list.

These basic concepts of list processing were first set forth in 1957 in a paper, [51], by A. Newell, J. C. Shaw and H. A. Simon. John McCarthy in 1960, [46], devised an important programming system and language, LISP, for list processing which permits *overlapping* of lists (a single cell can occur in several data lists) and which automatically reclaims most data lists which are no longer needed (the reclamation is called *garbage collection*). Collins in 1960, [10], devised an alternative scheme for overlapping using *reference counts*. A computer program system for infinite-precision integer arithmetic using list processing was first described by Collins in 1966, [11], although the system existed as early as 1961. We shall discuss list processing further in connection with polynomials. An in-depth treatment of list processing principles is to be found in the book [39] by D. E. Knuth.

Classical methods for base β arithmetic, as taught in the elementary schools for $\beta = 10$, are sufficiently well specified and efficient for computer algorithms, except for division. In division the method specified for generating the successive digits of the quotient involves some human judgement, which must be eliminated from a computer algorithm. In 1960, D. A. Pope and M. L. Stein, [54], proposed the following algorithm. We first normalize the divisor, multiplying both, it and the dividend, by a positive integer such that the leading digit of the new divisor is at least $\lceil \beta/2 \rceil$, the integral part of $\beta/2$. For this purpose we can use $\lceil \beta/(b_n + 1) \rceil$ as multiplier, where b_n is the leading digit of the original divisor. Now let $B = \sum_{i=0}^n b_i \beta^i$ be the normalized divisor, $Q = \sum_{i=0}^k q_i \beta^i$ the desired quotient. Assume q_k, \dots, q_{j+1} have already been determined and let $A = \sum_{i=0}^{n+j+1} a_i \beta^i$ be the current remainder. Then q_j is approximated by $\bar{q}_j = \lceil (a_{n+j+1}\beta + a_{n+j})/b_n \rceil$ unless $a_{n+j+1} = b_n$, in which case $\bar{q}_j = \beta - 1$. Pope and Stein showed that in all cases $0 \leq \bar{q}_j - q_j \leq 2$. Thus at most two corrections are required to obtain q_j and the necessity of a correction is always characterized by a negative remainder. A 1969 analysis by Collins and D. R. Musser [25], shows that, except for very small β , $\bar{q}_j - q_j$ has the values 0, 1 and 2 with probabilities of about .674, .318 and .008, respectively. D. E. Knuth, 1969 ([40], Section 4.3.1) has shown how to make a simple correction to \bar{q}_j , depending also on a_{n+j-1} and b_{n-1} , obtaining q_j^* such that $0 \leq q_j^* - q_j \leq 1$ and $q_j^* - q_j = 1$ with probability not exceeding $3/\beta$.

Any algorithm \mathcal{A} has a certain well-defined set $I_{\mathcal{A}}$ of valid inputs (the elements of $I_{\mathcal{A}}$ may be n -tuples). If the algorithm is initiated with a valid input x , the algorithm performs a certain finite number, $t_{\mathcal{A}}(x)$, of primitive actions (which may be thought of as the instructions of a real computer) and finally stops, producing

some output (which may be an m -tuple). $t_{\mathcal{A}}$ is the **computing time function** of the algorithm \mathcal{A} .

The computing time function of an algorithm obviously depends on the specific definition of a primitive action (i.e., on the particular computer on which the algorithm is implemented) and on various other uninteresting and inessential aspects of the algorithm, so we are interested only in the codominance equivalence class of the function in the following sense.

If f and g are two real-valued functions defined on a set S we say that f is **dominated by g** , or g **dominates f** , and write $f \leq g$ in case there is a positive real number c such that $f(x) \leq cg(x)$ for all $x \in S$. If $f \leq g$ and $g \leq f$, we say that f and g are **codominant**, and write $f \sim g$. If $f \leq g$ but not $g \leq f$ we say that f is **strictly dominated by g** and write $f < g$.

The **β -length** of a non-zero integer a , $L_{\beta}(a)$, is defined as the number of digits in its base β representation; the β -length of zero is defined to be 1. It is easy to verify that if γ is any other base, then $L_{\beta} \sim L_{\gamma}$ so we shall often just write $L(a)$ or refer to the length of a .

Now if M is the classical algorithm for multiplication of two infinite-precision integers, it is easy to verify, and intuitively obvious, that $t_M(a, b) \sim L(a) \cdot L(b)$ for $a \neq 0$ and $b \neq 0$. In particular, this means we can multiply any two n -digit numbers in at most n^2 units of time. It was not until 1962 that a faster integer multiplication algorithm was discovered, when A. Karatsuba proposed an algorithm with computing time dominated by $n \log_2 3$ ($\log_2 3 = 1.58 \dots$). Karatsuba's algorithm is the first in an infinite sequence of algorithms M_1, M_2, \dots , the computing time of M_k being dominated by n^{e_k} , where $e_k = \log_{k+1}(2k+1)$. Thus for every $\varepsilon > 0$ there is a multiplication algorithm with computing time dominated by $n^{1+\varepsilon}$. Still more recently, multiplication algorithms have been discovered by S. A. Cook and by A. Schonhage which are faster than any of the M_k . For an excellent description and analysis of these integer multiplication algorithms, see [40], Section 4.3.

Unfortunately, these "fast" multiplication algorithms are faster only for larger integers than normally arise in most algebraic calculations. James R. Pinkert found [52], that M_1 became faster than M only when the inputs had β -lengths greater than 55 (about 550 decimal digits). Subsequent discussions of computing times in this article will therefore be based on the use of classical algorithms for integer arithmetic.

Of course, the time to add or subtract two integers, a and b , is

$$\leq \max(L(a), L(b)) \sim L(a) + L(b).$$

The time to divide a by b using the classical algorithm, obtaining a quotient q and remainder r , is

$$\sim L(b) \cdot L(q) \sim L(b) \cdot (L(a) - L(b) + 1) \quad \text{for } |a| \geq |b| > 0.$$

The classical algorithm for computing the g.c.d. of two integers is the familiar

Euclidean algorithm. It can be shown, [15], that the time to compute $c = \gcd(a, b)$, $a \geq b > 0$, by the Euclidean algorithm is $\leq L(b) \cdot L(a/c) \sim L(b) \cdot (L(a) - L(c) + 1)$.

D. H. Lehmer [43], has devised a version of the Euclidean algorithm which is faster for large integers. Lehmer's version has the same computing time function as the ordinary Euclidean algorithm, to within codominance, but it is about ten times as fast on a typical computer when the inputs are longer than one β -digit. Knuth [41], has recently discovered an integer g.c.d. algorithm whose computing time, for inputs of length n , is dominated by $n^{1+\epsilon}$ for every $\epsilon > 0$. But this algorithm also appears impractical for integers of the sizes which commonly occur in algebraic calculations, and we shall not assume its use.

Instead of studying the computing time function of an algorithm directly, it is often more enlightening to study certain related functions. Consider, for example, the Euclidean algorithm E , and let $S(m, n, k)$ be the set of all pairs of integers (a, b) such that $|a| \geq |b| > 0$, $L(a) = m$, $L(b) = n$, and $L(c) = k$, where $c = \gcd(a, b)$. The sets $S(m, n, k)$ are finite, disjoint, and together contain all valid inputs to E . If

$$t_E^+(m, n, k) = \max\{t_E(a, b) : (a, b) \in S(m, n, k)\},$$

t_E^+ is the **maximum computing time function** for E . The result about t_E quoted above can be restated as $t_E^+(m, n, k) \leq n(m - k + 1)$. It can, in fact, be shown [15], that $t_E^+(m, n, k) \sim n(m - k + 1)$. Analogously, we can define a **minimum computing time function** t_E^- for E . It is proved in [15] that $t_E^-(m, n, k) \sim n(m - n + 1) + k(n - k + 1)$. Note that, therefore, $t_E^- < t_E^+$. An **average computing time function** for E is defined by

$$t_E^*(m, n, k) = \sum \{t_E(a, b) : (a, b) \in S(m, n, k)\} / \text{card}(S(m, n, k))$$

($\text{card}(S)$ is the cardinality of the set S). It is a much deeper result of [15], proved with the aid of a recent result of Dixon [27], that $t_E^* \sim t_E^+$. Dixon's result pertains to the average number of divisions in the Euclidean algorithm, a subject which is treated at length by Knuth in [40], Section 4.5.3.

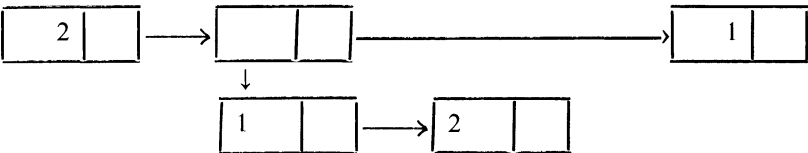
Among the various program systems which are now available for computer algebra, only a few provide infinite-precision integer arithmetic. Among these, one of the most readily available is SAC-1 since it is programmed, with the exception of a few simple "primitive" subprograms, entirely in the standardized Fortran IV language [1]. SAC-1 is organized in modular form, with different modules or "sub-systems" providing different capabilities. There are currently ten modules, of which the second [17], is for integer arithmetic while the first [16], provides a list, processing capability required by all other modules. The SAC-1 system is currently in use at some 50 institutions throughout the United States and in several other countries on computers of half a dozen major manufacturers. Other systems which provide infinite-precision integer arithmetic are Reduce 2 [31], PL/1-FORMAC (up to 2295 decimal digits!) [56], and SCRATCHPAD/1 [29]. Of these, only the former is available on non-IBM computers.

Codominance functions for the computing times are deliberately devoid of any constants and must be supplemented for practical purposes with empirically observed computing times for particular cases. Table 1 gives sample computing times for integer addition, multiplication, division and g.c.d. calculation. A , B and C are 100-decimal-digit integers, while D and E are 200 decimal digits long. Times, in seconds, are for the SAC-1 system on a UNIVAC 1108 computer. More extensive tables and formulas for empirical computing times are given in [17].

TABLE 1. Computing Times for Integer Arithmetic

$C = A \div B$.001
$D = A \cdot B$.007
$B = D / A$.018
$A = \text{gcd}(D, E)$.150

3. Polynomial arithmetic. We have seen above how to store in the computer any list (a_1, a_2, \dots, a_n) in which each a_i is a “small integer” (i.e., small enough to be stored in one memory location). We now have a need to consider more general lists. Let S be a set of small integers (e.g., $S = \{a: |a| < \beta\}$). A *list over S* is defined recursively as a sequence (a_1, \dots, a_n) such that, for $1 \leq i \leq n$, either $a_i \in S$ or a_i is a list over S . In this definition $n = 0$ is permitted, so that we have a null list over S , denoted by “()”. The *order* of a list over S may also be defined recursively: $\text{ord}(a) = 0$ for $a \in S$, $\text{ord}(()) = 1$ and, for $n > 0$, $\text{ord}((a_1, \dots, a_n)) = \max(\text{ord}(a_1), \dots, \text{ord}(a_n)) + 1$. For example, $(3, (), (2, (3, 1), 1), ((), 2))$ is a list of order 3. The *diagram* for the second order list $(2, (1, 2), 1)$ is



As suggested by the diagram, the vertical arrow emanating from the second word in the upper row contains in its element field the location of the list $(1, 2)$ shown in the lower row of the diagram. To preclude ambiguity, each word of a list may contain also a *type field* to designate whether its element field is a list location or an atom (i.e., element of S).

Now suppose that \mathcal{R} is some commutative ring for which a list representation has been specified. For each $a \in \mathcal{R}$, let \bar{a} be the list which represents a . Consider now the specification of a list representation for the polynomial ring $\mathcal{R}(x)$ of polynomials with coefficients in \mathcal{R} . The most obvious possibility is to represent any polynomial $A(x) = \sum_{i=0}^n a_i x^i$ with $a_n \neq 0$ by the list $(\bar{a}_n, \bar{a}_{n-1}, \dots, \bar{a}_0)$, and to represent the zero polynomial by the null list $()$. However, when working with **sparse polynomials**, i.e., polynomials with many zero coefficients, this representation

is wasteful of storage. Another possibility is then to express $A(x)$ in the form $\sum_{j=1}^k a_j x^{e_j}$, where $e_1 > e_2 > \dots > e_k$ and each $a_j \neq 0$, and to represent $A(x)$ by the list $(\bar{a}_1, e_1, \bar{a}_2, e_2, \dots, \bar{a}_k, e_k)$. Of course, for polynomials with few zero coefficients, this representation requires more memory than the former, but never more than twice as much.

Polynomials in several variables are conveniently regarded as polynomials in one **main variable** with coefficients which are polynomials in the remaining variables. The result is a **recursive representation** for $\mathcal{R}[x_1, \dots, x_r]$. It has the advantage that **recursive algorithms** can naturally be used for various operations — algorithms which use themselves as subalgorithms, directly or indirectly, to perform the same operation on polynomials in fewer variables. For operations such as addition or multiplication such a recursive algorithm is merely simpler, but for other operations such as polynomial division the only known algorithms depend on the use of a main variable, and are thus inherently recursive.

We denote by I , Q , R and C , respectively, the rings of the integers, the rationals, the real numbers and the complex numbers. For any polynomial $A(x_1, \dots, x_r)$ over C we define two “norms”, $|A|_\infty$ and $|A|_1$ by induction on r . For $r = 0$ ($A \in C$), we define $|A|_\infty = |A|_1 = |A|$. For $r > 0$, if $A(x_1, \dots, x_r) = \sum_{i=0}^n A_i(x_1, \dots, x_{r-1})x_r^i$, we define

$$(1) \quad |A|_\infty = \max_{0 \leq i \leq n} |A_i|_\infty$$

and

$$(2) \quad |A|_1 = \sum_{i=0}^n |A_i|_1.$$

$|A|_\infty$ and $|A|_1$ are called the **max norm** and **sum norm**, respectively. One obtains easily the relations

$$(3) \quad |A + B|_\infty \leq |A|_\infty + |B|_\infty,$$

$$(4) \quad |A + B|_1 \leq |A|_1 + |B|_1,$$

$$(5) \quad |A \cdot B|_\infty \leq |A|_\infty \cdot |B|_1,$$

$$(6) \quad |A \cdot B|_1 \leq |A|_1 \cdot |B|_1,$$

and

$$(7) \quad |A|_\infty \leq |A|_1.$$

If $A(x_1, \dots, x_r)$ is a polynomial in r variables, we denote by $\partial_i(A)$ the degree of A in x_i and by $\partial(A)$ the **degree vector** $(\partial_1(A), \dots, \partial_r(A))$. We shall often write $\deg(A)$ for $\partial_r(A)$, the degree of A in its main variable.

Let $m_i = \partial_i(A)$, $n_i = \partial_i(B)$, $a_0 = |A|_\infty$, $b_0 = |B|_\infty$, $a_1 = |A|_1$, $b_1 = |B|_1$. Clearly we can design an algorithm to add integral polynomials whose computing

time is dominated by

$$(8) \quad \{L(a_0) + L(b_0)\} \prod_{i=1}^r (m_i + n_i + 1).$$

A "classical" algorithm for integral polynomial multiplication has computing time dominated by

$$(9) \quad L(a_0)L(b_1) \prod_{i=1}^r (m_i + 1)(n_i + 1).$$

We can also design a "classical" algorithm for integral polynomial division with computing time dominated by (9) if we now let A be the divisor and B the resulting quotient. If we attempt to obtain a bound for the computing time of this algorithm depending only on the inputs, namely the divisor A and the dividend C , then we encounter the problem of determining an upper bound for $|B|_\infty$ in terms of $|A|_\infty$, $|C|_\infty$, $\partial(A)$ and $\partial(C)$, where $C = A \cdot B$. This problem appears to be no easier than the related problem of obtaining an upper bound for the norm of any divisor of A as a function of $|A|_1$ and $\partial(A)$. This latter problem is one which is presently far from a satisfactory solution even for the case $r = 1$, but we will consider some of the known results.

Define $U(m, n, d) = \max\{|B|_\infty : A(x), B(x) \in I[x] \& B|A \& \deg(A) = m \& \deg(B) = n \& |A|_1 = d\}$ for $m \geq n \geq 1$ and $d \geq 1$. It is easy to see that

$$(10) \quad U(m, 1, d) = U(m, m, d) = d.$$

Let $A(x), B(x) \in I[x]$, $B|A$, $\deg(A) = m$, $\deg(B) = n$ and $|A|_1 = d$. By considering the factorization of A over the field of complex numbers, it is not difficult to show that

$$(11) \quad |B|_1 \leq |b_n| \cdot (\bar{a} + 1)^n,$$

where $b_n = \text{lcf}(B)$, the leading coefficient of B and $\bar{a} = \max_{1 \leq i \leq m} |\alpha_i|$, $\alpha_1, \dots, \alpha_m$ being the zeros of A . We can also show quite easily that

$$(12) \quad \bar{a} \leq d - 1.$$

Since $|b_n| \leq |a_m| \leq d$, where $a_m = \text{lcf}(A)$, we have by (11) and (12) that

$$(13) \quad U(m, n, d) \leq d^{n+1}.$$

A second approach to this problem yields a different bound. If c is any integer, then $B(c)|A(c)$ in I . Hence if $A(c) \neq 0$, then $|B(c)| \leq |A(c)|$. Since A has at most m zeros, we can find $n + 1$ integers c such that $A(c) \neq 0$ and $|c| \leq [(m + n)/2]$, the least integer greater than or equal to $(m + n)/2$, and from the values of $B(c)$, the coefficients of c can be determined by interpolation. By considering the sum norms of the Lagrange interpolation polynomials, we obtain

$$(14) \quad U(m, n, d) \leq (m+1)^{m+n+1}d.$$

By elaboration of the interpolation approach, David R. Musser [50], obtains sharper bounds for $|B|_1$ under the assumption that $A(c) \neq 0$ for $|c| \leq [n/2]$, c an integer.

Can either of the bounds (13) and (14) be realized? A partial answer is obtained for the case $d = 2$ by taking $A(x) = x^m - 1$ and $B(x) = \Phi_m(x)$, the m th cyclotomic polynomial. Φ_m is an irreducible divisor of A of degree $\phi(m)$, where ϕ is Euler's function. Paul Erdős showed in 1945 [28], that for some $c_1 > 0$, $|\Phi_m|_\infty > \exp\{c_1(\log m)^{4/3}\}$ for infinitely many positive integers m . This implies that for every positive integer k there exist infinitely many pairs (m, n) for which

$$(15) \quad U(m, n, 2) > m^k.$$

Thus $U(m, n, d)$ is not bounded by any polynomial function of m, n and d , as one might otherwise conjecture.

Now let $A(x_1, \dots, x_r)$ be a multivariate integral polynomial, with $m_i = \partial_i(A)$. By induction on r , the interpolation approach to factor coefficient bounding can be used to prove that if $B|A$, then

$$(16) \quad |B|_1 \leq \left\{ \prod_{i=1}^r (m_i + 1)^{2m_i+1} \right\} |A|_1.$$

We shall see in Section 5 that a bound such as (16) is essential for an efficient algorithm for computing the complete factorization of an integral polynomial.

Thus far we have considered division in a polynomial ring $\mathcal{R}[x_1, \dots, x_r]$, with the implicit assumptions that \mathcal{R} is an integral domain and that the quotient is known to exist. If the quotient does exist then in an integral domain it is unique. If we drop the existence assumption then we have the problem of designing a **trial division algorithm** which, given A and $B \neq 0$, first decides whether $B|A$ and then, if so, computes $C = A/B$. By induction on r , we may assume a trial division algorithm for \mathcal{R} and obtain one for $\mathcal{R}[x]$, obtaining thereby a recursive algorithm for $\mathcal{R}[x_1, \dots, x_r]$. If $A = 0$, then $C = 0$. Otherwise, $B|A$ only if $m \geq n$ and $b|a$, where $m = \deg(A)$, $n = \deg(B)$, $a = \text{ldcf}(A)$ and $b = \text{ldcf}(B)$. If $m \geq n$ and $c = a/b$, then cx^{m-n} will be a term of the quotient, if it exists, and the process is repeated with $A_1(x) = A(x) - cx^{m-n}B(x)$ in place of $A(x)$.

In the case $\mathcal{R} = I$ there does indeed exist an obvious trial division algorithm, and so we obtain one for $I[x_1, \dots, x_r]$. This algorithm is efficient for cases in which the quotient exists, and one would expect it to terminate quickly in most cases for which the quotient does not exist. However, rigorous analysis of either worst case or average behavior appears to be very difficult. After a few iterations of leading coefficient division and subtraction, the coefficients of the remainder can become very large, with the result that the computing time can be very large. Such extreme cases can be easily constructed with $\text{ldcf}(B) = 1$, but other extreme cases are difficult to find. This is an excellent example of an algorithm whose behaviour is presumably satisfactory but not yet well understood.

In this section we have discussed some aspects of "classical" algorithms for performing the arithmetic operations in $\mathcal{R}[x_1, \dots, x_r]$, with emphasis on the case $\mathcal{R} = I$. Numerous computer program systems now exist with capabilities of this general nature, but with significant differences of importance to the user. The SAC-1 Polynomial System [18], provides polynomial arithmetic for $\mathcal{R} = I$ together with some other operations such as differentiation and substitution. The ALTRAN system [6] and [30], provides similar capabilities, but with limited-precision integer arithmetic. In ALTRAN, the polynomials may also have exact rational coefficients or approximate real coefficients. Also, ALTRAN is one of the few current systems which does not use either the recursive canonical form for polynomials or a list representation. Other widely distributed systems with comparable capabilities for polynomial arithmetic are Reduce 2 [31], and PL/I-FORMAC [57].

As an illustration of actual computing times, it requires approximately one second to multiply two univariate polynomials of degree 30 with coefficients 10 decimal digits long, using SAC-1 on a UNIVAC 1108 computer. The time to multiply two bivariate polynomials of degree 6 in each variable with coefficients 30 decimal digits long is approximately 5 seconds. The time to divide $A \cdot B$ by A is approximately the same as the time to multiply A by B .

4. Polynomial greatest common divisors. Computing g.c.d.'s of multivariate integral polynomials is very important because, apart from other applications, it permits us to perform arithmetic in the fraction field consisting of the rational functions over the field Q of rational numbers, while keeping fractions reduced to lowest terms.

There is a classical algorithm which is available for this purpose, namely the Euclidean algorithm for univariate polynomials over a field. Given $A(x_1, \dots, x_r)$, we can regard A as a polynomial in x_r with coefficients in $Q(x_1, \dots, x_{r-1})$. Assuming, inductively, a g.c.d. algorithm for $I[x_1, \dots, x_{r-1}]$, we can perform the required coefficient arithmetic in $Q(x_1, \dots, x_{r-1})$ and obtain a g.c.d. algorithm for $I[x_1, \dots, x_r]$. In the case $r = 1$ we have available the Euclidean algorithm for integers as a basis.

Let \mathcal{R} be a **g.c.d. domain**, that is, an integral domain in which any two elements have a g.c.d. If $a, b \in \mathcal{R}$, we say that a is an **associate** of b , and write $a \sim b$ in case there is a unit u such that $a = ub$. " \sim " is an equivalence relation. \mathcal{A} is called an **ample set** for \mathcal{R} in case \mathcal{A} has exactly one element in each equivalence class of associates. Relative to \mathcal{A} we can write $\gcd(a, b)$ for the unique g.c.d. of a and b in \mathcal{A} . If $A(x)$ is any non-zero polynomial over \mathcal{R} , $A(x) = \sum_{i=0}^n a_i x^i$, the **content** of A , written $\text{cont}(A)$, is $\gcd(a_0, a_1, \dots, a_n)$ and the **primitive part** of A , written $\text{pp}(A)$, is $A/\text{cont}(A)$. Thus $A = \text{cont}(A) \circ \text{pp}(A)$ and $\text{pp}(A)$ is *primitive*, that is, its content is a unit.

$\mathcal{R}[x]$ is also a g.c.d. domain, the units of $\mathcal{R}[x]$ are just the units of \mathcal{R} , and the polynomials in $\mathcal{R}[x]$ with leading coefficients in \mathcal{A} constitute an ample set for $\mathcal{R}[x]$. We have, for $A(x), B(x) \in \mathcal{R}[x]$, $A \neq 0$ and $B \neq 0$,

$$(17) \quad \gcd(A, B) \sim \gcd(\text{cont}(A), \text{cont}(B)) \cdot \gcd(\text{pp}(A), \text{pp}(B)).$$

In fact, if we choose an ample set which is **multiplicative** (closed under multiplication), then we shall have equality in (17). Therefore, it suffices to obtain a g.c.d. algorithm for primitive elements of $\mathcal{R}[x]$.

If A and B are non-zero elements of $\mathcal{R}[x]$, a **remainder** of A and B is a polynomial R such that for some polynomial Q and some $c, d \in \mathcal{R}$, $c \neq 0$ and $d \neq 0$,

$$(18) \quad cA = BQ + dR$$

and either $R = 0$ or $\deg(R) < \deg(B)$. If $\text{ldcf}(B)$ is a unit we always have a unique remainder with $c = d = 1$, called the **natural remainder** and denoted by $\text{rem}(A, B)$. In general, if $m = \deg(A)$, and $n = \deg(B)$, we may assume $m \geq n$ (otherwise A is a remainder with $Q = 0$ and $c = d = 1$) and then there is a unique remainder with $c = b^{m-n+1}$ and $d = 1$, $b = \text{ldcf}(B)$, called the **pseudo-remainder** and denoted by $\text{prem}(A, B)$.

A **polynomial remainder sequence** (p.r.s) is a sequence of polynomials $A_1, A_2, \dots, A_l, A_{l+1} = 0$ in which A_{i+2} is a remainder of A_i and A_{i+1} for $1 \leq i < l$. If A and B are nonzero polynomials, there always exists a p.r.s. with $A_1 = A$ and $A_2 = B$. The nonzero polynomials A and B are **similar**, written $A \approx B$, in case for some $c, d \in \mathcal{R}$, $c \neq 0$ and $d \neq 0$, $cA = dB$. If $A_1, A_2, \dots, A_l, A_{l+1} = 0$ is a p.r.s. for A and B , then $A_l \approx \gcd(A, B)$. Hence if A and B are primitive, then $\gcd(A, B) \sim \text{pp}(A_l)$. Since p.r.s.'s can be constructed in many different ways, this provides us with many possible g.c.d. algorithms, some of which are much better than others.

If \mathcal{R} is a field, we can use the **natural Euclidean algorithm**, in which $A_{i+2} = \text{rem}(A_i, A_{i+1})$. Even if \mathcal{R} is not a field (as for example $\mathcal{R} = I[x_1, \dots, x_{r-1}]$) we can generate the natural Euclidean p.r.s. over the fraction field of \mathcal{R} , multiply A_i by the least common multiple of its denominators to obtain $\tilde{A}_i \in \mathcal{R}[x]$, and compute $\text{pp}(\tilde{A}_i)$. It turns out that the natural Euclidean algorithm is a very bad algorithm because the numerators and denominators of the A_i grow very rapidly as i increases, and hence the computing time is large. The **monic Euclidean algorithm**, in which A_{i+2} is the monic polynomial which is similar to $\text{rem}(A_i, A_{i+1})$, is much better.

Let us consider the special but illustrative case in which $\mathcal{R} = I$, $\deg(A) = n$, $\deg(B) = n - 1$ and the p.r.s. $A_1, A_2, \dots, A_l, A_{l+1} = 0$ is **normal**, that is, $\deg(A_i) - \deg(A_{i+1}) = 1$ for $2 \leq i < l$. Suppose also that the coefficients of A and B are approximately d digits long. Then it can be shown that the numerator and denominator of A_i in the natural Euclidean p.r.s. have approximately $(i^2 - 3i + 3)d$ and $(i^2 - 3i + 2)d$ digits, respectively, for $3 \leq i \leq l$, and the computing time of the algorithm is codominant with $d^2 n^6 L(n)$, approximately. By contrast, in the monic Euclidean algorithm the numerator and denominator of A_i have approximately $2i - 3$ digits, and the computing time of the algorithm is codominant with $d^2 n^4 L(n)$, approximately. This illustrates the dramatic impact which can be made by a seemingly insignificant change in the p.r.s.

We can avoid an excursion into the fraction field by using instead the **primitive p.r.s. algorithm**, in which $A_{i+2} = \text{pp}(\text{prem}(A_i, A_{i+1}))$. In this algorithm, typically only about one fourth as many g.c.d.'s in \mathcal{R} are required as in the monic Euclidean algorithm. Since the g.c.d.'s in \mathcal{R} account for most of the computing time of either algorithm, we find that the primitive p.r.s. algorithm is several times faster, depending on \mathcal{R} , but the computing times of the two algorithms are codominant if $\mathcal{R} = I[x_1, \dots, x_r]$, $r \geq 0$.

We can avoid all g.c.d.'s in \mathcal{R} , except at the end in computing $\text{pp}(A_l)$, by setting $A_{i+2} = \text{prem}(A_i, A_{i+1})$, obtaining the **pseudo-remainder p.r.s. algorithm**. However, we find that in this case the length of the coefficients of A_i grows exponentially as a function of i ; as a result, the computing time of the algorithm is exponential in the degree n .

In 1966, G. Collins discovered, [12] and [13], that every term of any p.r.s. for A and B , is similar to some **subresultant** of A and B , if A and B are polynomials of positive degrees over any integral domain \mathcal{R} . Let $m = \deg(A)$, $n = \deg(B)$ and $0 \leq k < \min(m, n)$. Let M be the square matrix of order $m+n$ whose successive rows are the coefficients of $x^{n-1}A(x), \dots, xA(x), A(x), x^{m-1}B(x), \dots, xB(x), B(x)$. M is the **Sylvester matrix** of A and B , and $\det(M)$, the determinant of M , is the resultant of A and B . More generally, let M_k be the matrix whose rows are the coefficients of $x^{n-k-1}A(x), \dots, A(x), x^{m-k-1}B(x), \dots, B(x)$. M_k has $m+n-2k$ rows and $m+n-k$ columns. For $0 \leq j \leq k$, let $M_{k,j}$ be the square matrix consisting of the first $m+n-2k-1$ columns of M_k followed by column $m+n-k-j$. Then $S_k(x) = \sum_{j=0}^k \det(M_{k,j})x^j$ is the k th subresultant of A and B , a polynomial of degree k or less.

By 1968 Collins, and independently W. S. Brown, had proved the *fundamental theorem of p.r.s.'s*, which gives explicit formulas relating the terms of any p.r.s. to their similar subresultants. Let $A_1, A_2, \dots, A_l, A_{l+1}$ be a p.r.s. of A and B defined by

$$(19) \quad e_i A_i = A_{i+1} Q_i + f_i A_{i+2}$$

for $1 \leq i \leq l-1$ (f_{l-1} is arbitrary). Let $c_i = \text{lcf}(A_i)$, $n_i = \deg(A_i)$ and $\delta_i = n_i - n_{i+1}$. The fundamental theorem states that

$$(20) \quad \left\{ \prod_{i=2}^{k-1} (-1)^{(n_{i-1}-n_k)(n_i-n_k)} f_{i-1}^{n_i-n_k} c_i^{\delta_{i-1}+\delta_i} \right\} c_k^{\delta_{k-1}-1} A_k$$

$$= \left\{ \prod_{i=2}^{k-1} e_{i-1}^{n_i-n_k} \right\} S_{n_k} \quad (3 \leq k \leq l),$$

$$(21) \quad \left\{ \prod_{i=2}^{k-1} (-1)^{(n_{i-1}-n_{k-1}+1)(n_i-n_{k-1}+1)} f_{i-1}^{n_i-n_{k-1}+1} c_i^{\delta_{i-1}+\delta_i} \right\} A_k$$

$$= \left\{ \prod_{i=2}^{k-1} e_{i-1}^{n_i-n_{k-1}+1} \right\} c_{k-1}^{\delta_{k-1}-1} S_{n_{k-1}-1} \quad (3 \leq k \leq l),$$

$$(22) \quad S_j = 0 \text{ for } n_k < j < n_{k-1} - 1 \quad (3 \leq k \leq l),$$

and

$$(23) \quad S_j = 0 \text{ for } 0 \leq j < n_l.$$

A proof of the fundamental theorem appears in [7].

The coefficients of the subresultants are determinants of known order, so the degrees and coefficient sizes of these coefficients can be estimated in terms of the degrees and coefficient sizes of the coefficients of A and B . Using the fundamental theorem, one can then estimate degrees and coefficient sizes for a specified p.r.s., and from this the computing time of a g.c.d. algorithm based on this p.r.s. All of the results stated or referred to above were obtained in this way and it is difficult to imagine that they could have been derived otherwise.

Two new types of p.r.s.'s were introduced by Collins in [13] which avoid g.c.d. calculations in \mathcal{R} but which, unlike the pseudo-remainder p.r.s., control coefficient growth in most cases. The **reduced p.r.s.** is defined by setting $e_i = c_{i+1}^{\delta_i+1}$ for $1 \leq i < l$, $f_1 = 1$ and $f_i = e_{i-1}$ for $2 \leq i < l-1$ in (19). The **subresultant p.r.s.** is defined by $A_i = S_{n_i-1}$, for $3 \leq i \leq l+1$, but a process was obtained for computing A_{i+2} from $\text{prem}(A_i, A_{i+1})$ and previous terms of the sequence. In either p.r.s., $\text{prem}(A_i, A_{i+1})$ is divided by certain powers of the c_j for $j \leq i$ to obtain A_{i+2} , but the coefficients of the A_i remain in \mathcal{R} for any integral domain \mathcal{R} . This is true by definition for the subresultant p.r.s. and by the fundamental theorem we have

$$(24) \quad A_k = \left\{ \prod_{i=2}^{k-2} (-1)^{(n_{i-1}-n_{k-1}+1)(n_i-n_{k-1}+1)} c_i^{\delta_{i-1}(\delta_i-1)} \right\} (-1)^{\delta_{k-2}-1} S_{n_{k-1}-1} \quad (3 \leq k \leq l),$$

where $A_1, A_2, \dots, A_l, A_{l+1} = 0$ is the reduced p.r.s. and $c_i = \text{ldcf}(A_i)$. Each exponent $\delta_{i-1}(\delta_i-1)$ of c_i is non-negative and indeed, each exponent is zero just in case the p.r.s. is normal, in which case the two p.r.s.'s differ only in signs. For randomly chosen polynomials A and B the p.r.s. is almost always normal as observed by Knuth ([40], Sec. 4.6.1). In such normal cases coefficient growth is controlled by the reduced p.r.s. algorithm; its computing time is codominant with that of the primitive p.r.s. algorithm, but typically several times faster. However, in severely non-normal cases coefficient growth and computing time for the reduced p.r.s. algorithm can be exponential in the degree n . The subresultant p.r.s. algorithm controls coefficient growth in these non-normal cases, but it performs so many multiplications and divisions in doing so that its computing time is also exponential.

The foregoing developments relating to the theory of polynomial remainder sequences and subresultants did much for polynomial g.c.d. calculation, but the problem remained in a somewhat unsatisfactory state prior to the introduction of modular algorithms in 1968.

In order to focus on the essential ideas, let us consider first just the case $\mathcal{R} = I$, and later generalize to $\mathcal{R} = I[x_1, \dots, x_r]$. There is a unique homomorphism, ϕ_p , of I onto the finite field $GF(p) = \{0, 1, \dots, p-1\}$ of the integers modulo p such that $\phi_p(1) = 1$, for any prime p ; ϕ_p is called a **modular homomorphism**. It is easy to see that if ϕ is any homomorphism from any g.c.d. domain, \mathcal{R} , to another, S , then $\phi(\gcd(A, B)) \mid \gcd(\phi(A), \phi(B))$ for any polynomials A and B over \mathcal{R} . Hence if $\phi(A) \neq 0$ or $\phi(B) \neq 0$, then $\deg(\gcd(\phi(A), \phi(B))) \geq \deg(\phi(\gcd(A, B)))$. If $\deg(\gcd(\phi(A), \phi(B))) = \deg(\phi(\gcd(A, B)))$, then $\gcd(\phi(A), \phi(B)) \approx \phi(\gcd(A, B))$. Let $C = \gcd(A, B)$ and $k = \deg(C)$. Assume $\phi(\text{ldcf}(A)) \neq 0$ and $\phi(\text{ldcf}(B)) \neq 0$. Then $\phi(S_k(A, B)) = S_k(\phi(A), \phi(B))$, where $S_k(A, B)$ denotes the k th subresultant of A and B . By the fundamental theorem, then, $\deg(\gcd(\phi(A), \phi(B))) = k$ if and only if $\deg(\phi(S_k(A, B))) = k$, and this is equivalent to the condition that $\phi(\det(M_{k,k})) \neq 0$, since $\det(M_{k,k})$ is the leading coefficient of $S_k(A, B)$.

Thus in the case $\mathcal{R} = I$, $\gcd(\phi_p(A), \phi_p(B)) \approx \phi_p(\gcd(A, B))$ unless p divides $\text{ldcf}(A)$, $\text{ldcf}(B)$, or $\det(M_{k,k})$. There are clearly only a finite number of such **unlucky primes**. We also have, for any \mathcal{R} , $\text{ldcf}(\gcd(A, B)) \mid \gcd(\text{ldcf}(A), \text{ldcf}(B))$. Thus there is a polynomial \bar{C} similar to C whose leading coefficient is $\bar{c} = \gcd(a, b)$ where $a = \text{ldcf}(A)$ and $b = \text{ldcf}(B)$. For the field $GF(p)$ we can use $\{0, 1\}$ as ample set, so that $C_p^* = \gcd(\phi_p(A), \phi_p(B))$ is monic. Multiplying C_p^* by $\bar{c}^* = \phi_p(\bar{c})$, we obtain \bar{C}_p^* and $\bar{C}_p^* = \phi_p(\bar{C})$ provided p is lucky. From a sufficient number of such \bar{C}_p^* we can compute \bar{C} using an algorithm for the Chinese remainder theorem, and then $C = \text{pp}(\bar{C})$, assuming as before that A and B are primitive.

In following this plan, one uses a precomputed list of distinct primes, each nearly as large as can be stored in one computer memory location. An a priori bound U is computed for $\det(M_{k,k})$ and the coefficients of \bar{C} . Primes which divide a or b , or which produce non-minimal g.c.d. degrees relative to other primes, are rejected. When primes p_1, \dots, p_r with $\prod_{i=1}^r p_i > 2U$ have been used and retained, the Chinese remainder theorem is applied. This algorithm has a computing time dominated by $n^3 L(d)^2$, where the inputs have degrees of n or less and sum norms of at most d .

The algorithm just described ordinarily uses far more primes than actually necessary. In place of this a priori bound method one may use a trial division method, together with an iterative algorithm for the Chinese remainder theorem. After processing each prime, the Chinese remainder theorem is applied to incorporate the \bar{C}_p^* just computed. If the result \tilde{C} so obtained is unchanged from the previous application, trial divisions of $\bar{c}A$ and $\bar{c}B$ by \tilde{C} are performed, using a modular algorithm. If \tilde{C} is a common divisor of $\bar{c}A$ and $\bar{c}B$, then $\tilde{C} = \bar{C}$ and $C = \text{pp}(\tilde{C})$; otherwise another prime is processed. If $\deg(\bar{C}_p^*) < \deg(\tilde{C})$ then \tilde{C} is discarded and a new \tilde{C} is formed from \bar{C}_p^* . The trial divisions are themselves performed by a modular algorithm. Provided that only a negligible number of unlucky primes are processed, and provided that C , A/C and B/C have sum norms of d or less, the computing time of this g.c.d. algorithm is dominated by $n^2 L(d) + nL(d)^2$. One may argue intuitively

that this assumption will almost always be satisfied and that accordingly the average computing time of the algorithm is codominant with $n^2L(d) + nL(d)^2$. But the maximum computing time of the algorithm may be as large as $n^3L(d)^2$. The algorithm is further enhanced by designing it to terminate with $C = 1$ whenever a prime p is found for which $C_p^* = 1$. If $C = 1$ then $C_p^* = 1$ for every lucky prime p , so one may likewise argue intuitively that the average computing time of the algorithm is codominant with $n^2 + nL(d)$ for relatively prime inputs.

The method just described generalizes readily to produce a g.c.d. algorithm for $I[x_1, \dots, x_r]$, assuming the availability of a g.c.d. algorithm for $\text{GF}(p)[x_1, \dots, x_r]$. The obvious method is to regard $A(x_1, \dots, x_r) \in I[x_1, \dots, x_r]$ as a polynomial in x_r with coefficients in $I[x_1, \dots, x_{r-1}]$. One must then begin by computing the primitive parts of A and B , so a g.c.d. algorithm is also required for $I[x_1, \dots, x_{r-1}]$. Brown observed [5], however, that a more efficient algorithm is obtained by treating the variables symmetrically. Define the **integer content** of A , $\text{cont}_I(A)$, as the g.c.d. of the integer coefficients of A and define the **integer primitive part** of A by $\text{pp}_I(A) = A/\text{cont}_I(A)$. Then the role of (17) is played by the formula

$$(25) \quad \text{gcd}(A, B) \sim \text{gcd}(\text{cont}_I(A), \text{cont}_I(B)) \cdot \text{gcd}(\text{pp}_I(A), \text{pp}_I(B)).$$

And in place of rejecting primes which produce g.c.d.'s of non-maximal degree in x_r , we reject all those which produce g.c.d.'s with non-maximal degree vectors, relative to some lexicographical ordering. Furthermore, we replace the leading coefficient of A by the **integer leading coefficient** of A , $\text{ldcf}_I(A)$, and set $\bar{c} = \text{gcd}(\text{ldcf}_I(A), \text{ldcf}_I(B))$. Using these replacements, Brown's method avoids some g.c.d. calculations in $I[x_1, \dots, x_{r-1}]$ at a cost of numerous g.c.d. calculations in I . The codominance class of the computing time is not affected, but experiments conducted by the author showed a substantial speed advantage for Brown's method.

A completely analogous method may be used to obtain a g.c.d. algorithm for $\text{GF}(p)[x_1, \dots, x_r]$, given one for $\text{GF}(p)[x_1, \dots, x_{r-1}]$ whenever $r \geq 2$, with evaluation homomorphisms replacing modular homomorphisms. The evaluation homomorphism ψ_a , for $a \in \text{GF}(p)$, is defined by $\psi_a(A(x_r)) = A(a)$. For evaluation homomorphisms, the analogue of the Chinese remainder is interpolation, and degree bounds replace integer coefficient bounds. Actually, an evaluation homomorphism is a special type of modular homomorphism, in which the modulus is the irreducible polynomial $x_r - a$, and interpolation is a special form of the Chinese remainder theorem; see [44] for an excellent comprehensive treatment of the Chinese remainder theorem and interpolation. Brown's symmetric treatment is again available, for the variables x_2, \dots, x_r , with $\text{GF}(p)[x_1]$ taking the place of I , and the trial divisions are themselves performed using evaluation homomorphisms and interpolation.

Brown has further improved the algorithm just described by eliminating the trial divisions. Let $A_1 = A/C$ and $B_1 = B/C$; A_1 and B_1 are called the **cofactor polynomials** of A and B . A_1 and B_1 are similar to polynomials \bar{A}_1 and \bar{B}_1 with $\text{ldcf}_I(\bar{A}_1) = a$ and $\text{ldcf}_I(\bar{B}_1) = b$, and we have $\bar{A}_1\bar{C} = \bar{c}A$, $\bar{B}_1\bar{C} = \bar{c}B$. Hence whenever p is lucky

we have $\phi_p(\bar{A}_1) = \phi_p(\bar{c})\phi_p(A)/\bar{C}_p^* = \bar{A}_p^*$ and $\phi_p(\bar{B}_1) = \phi_p(\bar{c})\phi_p(B)/\bar{C}_p^* = \bar{B}_p^*$. By applying the Chinese remainder theorem to the \bar{A}_p^* and \bar{B}_p^* , \bar{A}_1 and \bar{B}_1 can be computed for which $\bar{A}_1\bar{C} \equiv \bar{c}A$ (modulo M) and $\bar{B}_1\bar{C} \equiv \bar{c}B$ (modulo M), where M is the product of all primes used to obtain \bar{C} via the Chinese remainder theorem. Hence if $M > 2 \max\{|\bar{c}A|_1, |\bar{A}_1|_1 |\bar{C}|_1\}$ then $\bar{A}_1\bar{C} = \bar{c}A$, and a test of this inequality replaces trial division of $\bar{c}A$ by \bar{C} . The cofactor polynomials are frequently a useful byproduct of the g.c.d. calculation; for example, if A and B are the numerator and denominator of a rational function, then A_1 and B_1 are the numerator and denominator after reduction to lowest terms.

If the degrees of A and B are at most n in each of the r variables, and if the sum norms of A , B , C , A_1 and B_1 are all at most d , and if a negligible number of unlucky homomorphisms are encountered, then, as follows from the discussion in [5], the computing time of this modular g.c.d. algorithm is dominated by $n^{r+1}L(d) + n^rL(d)^2$. Hence one may conjecture that the average computing time of the algorithm is co-dominant with $n^{r+1}L(d) + n^rL(d)^2$. This is the same as the time required to multiply A by B , using a modular algorithm, and much less than the time to multiply A by B , using the classical algorithm.

The SAC-1 subsystem [20] provides modular algorithms for g.c.d. calculation, resultant calculation, multiplication, division and trial division in $I[x_1, \dots, x_r]$. As an example of actual computing times, let A , B and C be pairwise relatively prime univariate integral polynomials of degree 40 with coefficients which are eight decimal digits long. Table 2 gives computing times in seconds for the SAC-1 system on a UNIVAC 1108 computer for various operations and algorithms. More extensive tables are given in [20].

TABLE 2. Polynomial Algebra Computing Times

$\gcd(A, B)$.41
$\gcd(AC, BC)$	3.60
$A \cdot C$ (modular)	1.06
$A \cdot C$ (classical)	1.81
AC/C (modular)	1.16
AC/C (classical)	4.27

5. Polynomial factorization. The spectacular progress of recent years in obtaining efficient algorithms for polynomial g.c.d. calculation has been matched by equally significant progress in obtaining efficient polynomial factorization algorithms. As one might easily guess, the factorization problem is much more difficult than the g.c.d. problem, and so the current state of development of factorization algorithms is less advanced, but the achievements of the last five years have been remarkable!

If \mathfrak{A} is any unique factorization domain (u.f.d.) and a is any non-zero element of \mathfrak{A} , a can be expressed in the form $a = u \prod_{i=1}^r a_i^{e_i}$ where u is a unit, the a_i are irreducible, and the e_i are positive integers. Such an expression is unique to within

associates and the order of the irreducible factors. We will assume that we are given a multiplicative ample set \mathcal{A} for \mathfrak{U} . If we then require that the $a_i \in \mathcal{A}$ and assume also that $a \in \mathcal{A}$ then we must have $u = 1$ and the set of pairs (a_i, e_i) is unique. We may then refer to $\{(a_1, e_1), \dots, (a_r, e_r)\}$ as the **complete factorization** of a . Also, we are interested in the case that \mathfrak{U} is a polynomial domain $\mathfrak{U}_0[x]$ and \mathfrak{U}_0 has a multiplicative ample set \mathcal{A}_0 . We will then always choose as ample set for \mathfrak{U} the set \mathcal{A} consisting of all polynomials over \mathfrak{U}_0 with leading coefficients in \mathcal{A}_0 . Also, we shall be concerned primarily with the cases $\mathfrak{U} = \text{GF}(q)[x_1, \dots, x_s]$ and $\mathfrak{U} = I[x_1, \dots, x_s]$ although several of the algorithms to be discussed are of more general applicability.

The "classical" algorithm for polynomial factorization is Kronecker's method, which goes back at least to 1882 [42]. Kronecker's method is applicable whenever the coefficient domain \mathfrak{U}_0 is infinite and has only a finite number of units, and provides us with a complete factorization algorithm for $\mathfrak{U}_0[x]$ whenever we already have one for \mathfrak{U}_0 . In particular, it provides a complete factorization algorithm for $I[x_1, \dots, x_s]$ by induction on s .

For every positive integer k , Kronecker's algorithm enables us to find all factors of degree k of a given primitive polynomial A . From any such algorithm we can obtain an algorithm for the complete factorization of any polynomial A , as follows.

1. Set $B = \text{pp}(A)$ and set $k = 1$.
2. Find all factors of B of degree k .
3. Divide B as often as possible by the factors found in step 2, and add 1 to k .
4. If $k \leq \deg(B)/2$, go back to step 2.

Whenever step 2 is performed, the polynomial B has no factors of degree less than k , and so the irreducible factors of $\text{pp}(A)$ are all the factors found in step 2 together with the polynomial B (unless B is a unit) after step 4 is last performed. The multiplicity of each factor is discovered in step 3, and the content of A is factored separately in \mathfrak{U}_0 .

If A and B are polynomials over \mathfrak{U}_0 and $B|A$, then $B(a)|A(a)$ for every $a \in \mathfrak{U}_0$. If $A(a) \neq 0$, then $A(a)$ has only a finite number of divisors since \mathfrak{U}_0 has only a finite number of units. Since \mathfrak{U}_0 is infinite we can find $k+1$ distinct elements of \mathfrak{U}_0 , say a_0, a_1, \dots, a_k , such that $A(a_i) \neq 0$ for all i . If F_i is the finite set of factors of $A(a_i)$ and $\overline{\mathfrak{U}}_0$ is the fraction field of \mathfrak{U}_0 , we can compute by interpolation the set of all polynomials B over $\overline{\mathfrak{U}}_0$ of degree k or less such that $B(a_i) \in F_i$ for $0 \leq i \leq k$. After discarding those polynomials B of degree less than k and those with coefficients not in \mathfrak{U}_0 , Kronecker's algorithm continues by performing a trial division of A by each remaining polynomial.

Unless \mathfrak{U}_0 has characteristic zero it has at least two units, each F_i has at least two members, and the number of interpolations performed, for given k , is at least 2^{k+1} . If A is irreducible and of degree n , then k will assume all values up to $[n/2]$, so the maximum computing time of Kronecker's algorithm is an exponential function of n . In case $\mathfrak{U}_0 = I$, the maximum computing time of the algorithm for polynomials

with norms less than d will also be an exponential function of $L(d)$, for there is no known algorithm for complete factorization in I whose computing time is dominated by a polynomial function of the length of the integer to be factored. Numerous devices have been proposed for making Kronecker's algorithm more efficient (see, e.g., [38]) in the case $\mathfrak{U}_0 = I$, but these devices do not eliminate the exponential character of the algorithm, nor do they succeed in making the algorithm practically useful.

Kronecker's algorithm is also applicable when $\mathfrak{U}_0 = \text{GF}(q)[x_1, \dots, x_s]$, but again not very practical. It is not applicable when $\mathfrak{U}_0 = \text{GF}(q)$ and this is the case, as one might guess by analogy with the g.c.d. problem, which is crucial for the other cases. No general and reasonably efficient algorithm was known for the complete factorization of univariate polynomials over a finite field until 1967, when such an algorithm was discovered by Elwyn R. Berlekamp. Since Berlekamp's algorithm ([2]; see also [3], Section 6.1, and [40], Section 4.6.2) is applicable only to squarefree polynomials, we shall first consider how the reduction of the problem to this case can be achieved.

If $\prod_{i=1}^r A_i^{e_i}$ is the complete factorization of A in any u.f.d. \mathfrak{U} , then A is **squarefree** in case each $e_i = 1$. The product $\bar{A} = \prod_{i=1}^r A_i$ is the **greatest squarefree divisor** of A , $\bar{A} = \text{gsfd}(A)$. Let $k = \max_{1 \leq i \leq r} e_i$ and define $B_j = \prod_{e_i=j} A_i$ for $1 \leq j \leq k$. Then

$$(26) \quad A = \prod_{j=1}^k B_j^j$$

is the **squarefree factorization** of A . The B_j are squarefree and pairwise relatively prime, and $B_k \neq 1$. Conversely, these conditions completely characterize the B_j .

Now assume that \mathfrak{U} is a polynomial domain, and denote by A' the derivative of A . If \mathfrak{U} has characteristic zero then $\text{gcd}(A, A') = \prod_{i=1}^r A_i^{e_i-1} = \prod_{j=2}^k B_j^{j-1}$, and so

$$(27) \quad \text{gsfd}(A) = A / \text{gcd}(A, A').$$

From the complete factorization of $\bar{A} = \text{gsfd}(A)$, we can obtain the complete factorization of A by repeated trial division since A and \bar{A} have the same irreducible divisors. However, it will generally be advantageous to first compute the squarefree factorization and then compute the complete factorization of each B_j which is not unity.

For a domain of arbitrary characteristic, define

$$(28) \quad \delta_i = \begin{cases} 0, & \text{if } (A_i^{e_i})' = 0 \\ 1, & \text{otherwise} \end{cases},$$

$$(29) \quad \bar{B}_j = \prod_{e_i=j \text{ \& } \delta_i=1} A_i,$$

and

$$(30) \quad S = \prod_{\delta_i=0} A_i^{e_i}.$$

Then

$$(31) \quad A = \left\{ \prod_{j=1}^k \bar{B}_j^j \right\} S,$$

and Musser [50], has given an algorithm which computes the \bar{B}_j and S . If $S = 1$, in particular whenever the characteristic is zero, we obtain the squarefree factorization. Let $C_i = \{\prod_{j=i+1}^k \bar{B}_j^{-i}\}S$ and $D_i = \prod_{j=i}^k B_j$ for $1 \leq i \leq k+1$. Musser's algorithm computes $C_1 = \gcd(A, A')$, $D_1 = A/C_1$, $D_{i+1} = \gcd(C_i, D_i)$, $C_{i+1} = C_i/D_{i+1}$, $B_i = D_i/D_{i+1}$, and terminates when $D_i = 1$ with $i = k+1$ and $C_k = S$.

If $S \neq 1$ and \mathfrak{A} is a domain of multivariate polynomials, say $\mathfrak{A}_0[x_1, \dots, x_s]$, Musser's algorithm can be applied to S with some variable x_i in place of x such that $\partial S / \partial x_i \neq 0$ and a further factorization of S will result. If finally some $S \neq 1$ is obtained with $\partial S / \partial x_i = 0$ for all i , and if \mathfrak{A}_0 is a finite field, $\text{GF}(p^n)$, then Musser observes that S is the p th power of some polynomial T . We shall have $S(x_1, \dots, x_s) = \sum_i a_i x_1^{e_{i1}} \dots x_s^{e_{is}}$ and then $T(x_1, \dots, x_s) = \sum_i a_i^{1/p} x_1^{e_{i1}} \dots x_s^{e_{is}}$ where $a_i^{1/p} = a_i^{p^{n-1}}$.

Additions and subtractions can be performed in $\text{GF}(q)$ in time dominated by $L(q)$, multiplications and divisions in $L(q)^2$. If $a \in \text{GF}(q)$, a^k can be computed using at most $2 \log_2 k$ multiplications (see [40], Section 4.6.3), so the time to compute $a^{1/p}$ in $\text{GF}(q)$ is dominated by $L(q)^3$.

The input to Berlekamp's 1967 algorithm is a monic squarefree univariate polynomial A over $\text{GF}(q)$. The algorithm has two phases. The first phase determines the number, r , of irreducible monic factors of A , and has a computing time dominated by $n^3 L(q)^2 + n^2 L(q)^3$. If $r = 1$, then A is irreducible and the algorithm terminates. Otherwise, the second phase must be performed, for which the computing time is dominated by $n^2 r q L(q)^2$.

Berlekamp's algorithm therefore provides an efficient irreducibility test for univariate polynomials over $\text{GF}(q)$. If A is a squarefree univariate integral polynomial then $\text{discr}(A)$, the discriminant of A , is a non-zero integer. If p is a prime which is not a divisor of $\text{ldcf}(A)$, then $\text{discr}(\phi_p(A)) = \phi_p(\text{discr}(A))$, so if p does not divide $\text{discr}(A)$ then $\text{discr}(\phi_p(A)) \neq 0$ and $\phi_p(A)$ is squarefree. Hence $\phi_p(A)$ is squarefree for all but a finite number of primes, and we can readily find a prime for which $\phi_p(A)$ is squarefree; in fact, for given p , $\phi_p(A)$ is squarefree with probability $1 - 1/p$ ([40], Section 4.6.2). If $\phi_p(A)$ is irreducible then so is A , and we may seek to prove the irreducibility of integral polynomials in this manner. However, if $\deg(A) = n$ the probability that $\phi_p(A)$ is irreducible is only about $1/n$ ([40], Section 4.6.2), so the average number of trials required will be about n . Moreover, one cannot decide irreducibility in this manner for there are irreducible polynomials A such that

$\phi_p(A)$ is reducible for every prime p . A simple example is the polynomial $A(x) = x^4 + 1$ ([40], Section 4.6.2).

One might consider a complete factorization algorithm for primitive squarefree univariate integral polynomials based on the Chinese remainder theorem. After computing a bound B on the coefficients of the divisors of A , one would factor $\phi_p(A)$ for primes p_1, \dots, p_k such that $\prod_{i=1}^k p_i > 2B$. However, the time to apply the second phase of Berlekamp's algorithm to all of the $\phi_{p_i}(A)$ may be proportional to B , and B itself may be an exponential function of n as observed in the previous section. Moreover, if A has two irreducible factors, A_1 and A_2 , of the same degree, there is no apparent way to distinguish $\phi_p(A_1)$ from $\phi_p(A_2)$ in applying the Chinese remainder theorem, so it will have to be applied in all of the 2^{k-1} possible ways.

A superior alternative to the Chinese remainder theorem is provided by Hensel's p -adic lemma, [35], as proposed by Hans Zassenhaus in 1968, [57]. If $\phi_p(A)$ is squarefree, from the complete factorization of A modulo p^j we obtain by Hensel's algorithm the complete factorization of A modulo p^{j+1} . Starting with $j = 1$, we eventually obtain the complete factorization of A modulo $m = p^k$ with $p^k > 2B$. If A has r irreducible monic factors modulo p , then A also has exactly r irreducible monic factors modulo $m = p^k$. Thus $A \equiv aA_1 \cdots A_r \pmod{m}$, where $a = \text{lcf}(A)$ and the A_i are monic integral polynomials, irreducible modulo m . If B is an irreducible factor of A over the integers, then B is similar to a polynomial \bar{B} with $\text{lcf}(\bar{B}) = a$. \bar{B} must then be the product, modulo m , of a and some subset of $S = \{A_1, \dots, A_r\}$, and $B = \text{pp}(\bar{B})$. By considering smallest subsets of S first, the irreducible factors of A over I can be obtained after at most 2^{r-1} trial divisions.

Musser [50], has shown that the time for application of Hensel's algorithm is dominated by $n^2 L(m)^3 + nL(m)^2 L(c)$ where $n = \deg(A)$ and $c = |A|_1$. In fact, there is a "quadratic" version of Hensel's algorithm which proceeds from a factorization modulo p^j to one modulo p^{2j} , and Musser has shown that the time for this version is dominated by $n^2 L(m)^2 + nL(m)L(c)$. Since $L(m) \leq nL(c)$ for one of the bounds considered in Section 3, the time to obtain the factorization modulo m can be dominated by $n^4 L(c)^2$.

If A is irreducible, but $\phi_p(A)$ splits into linear factors in $\text{GF}(p)$, then in fact $2^{n-1} - 1$ subsets of S will have to be processed and the computing time of the algorithm will be exponential in n . It would therefore seem advisable to factor $\phi_p(A)$ for several small primes for which $\phi_p(A)$ is squarefree and then apply Hensel's lemma for a p which produces the smallest number of irreducible factors. But however many primes are used, the maximum computing time of the resulting algorithm will still be exponential in n , for H. P. F. Swinnerton-Dyer has shown (see [4]) that there are irreducible integral polynomials of degree n which have at least $n/2$ irreducible factors modulo p for every prime p . By considering the norms of these Swinnerton-Dyer polynomials Musser (unpublished paper) has shown that the maximum computing time of any "Berlekamp-Hensel" algorithm is not dominated

by any polynomial function of $n = \deg(A)$ and $c = |A|_1$. Musser also reestablishes this by consideration of cyclotomic polynomials.

Nevertheless, the average computing time of the Berlekamp-Hensel algorithm in which one uses the first p for which $\phi_p(A)$ is squarefree may have an average computing time which is dominated by a polynomial function of n and $L(c)$. For, if $A_{n,p}$ is the average number of irreducible monic factors of $\phi_p(A)$ for a random polynomial A of degree n , then $\lim_{p \rightarrow \infty} A_{n,p} = H_n = \sum_{i=1}^n 1/i$ (see [40], Section 4.6.2, or [3], Chapter 3), and $H_n \leq \ln(n) + 1$. So the number of modulo m factors which must be considered for an "average" polynomial of degree n does not exceed $2^{\ln n} = n^{\ln 2} = n^{0.693\cdots} < n$.

A version of the Berlekamp-Hensel algorithm with some additional improvements has been detailed, implemented as part of the SAC-1 system, and subjected to experimentation by Musser [50] and [26]. Musser's algorithm contains a parameter v , and for each of v primes p for which $\phi_p(A)$ is squarefree, it computes the degree set of $\phi_p(A)$, which is the set consisting of the degrees of all factors of $\phi_p(A)$. The degree set of A must be contained in the intersection of the degree sets of the $\phi_p(A)$, and this is used to reduce the number of modulo m factors which must be tried. From among these v primes, Hensel's algorithm is applied to one for which the number r of irreducible factors is least (unless $r = 1$).

The degree set of $\phi_p(A)$ is obtained not by use of Berlekamp's algorithm, but by use of a **distinct degree factorization** algorithm described in [40]. Given a monic squarefree polynomial D over $\text{GF}(q)$, this algorithm computes polynomials B_1, \dots, B_k such that $B = \prod_{i=1}^k B_i$, and B_i is the product of all monic irreducible divisors of B of degree i . This algorithm has a computing time dominated by $n^3 L(q)^2 + n^2 L(q)^3$; this is the same time as the first phase of Berlekamp's algorithm, which only determines the number of irreducible divisors, while the degree set is determined by the distinct degree factorization.

Musser's algorithm further reduces the number of modulo m factors which must be computed, by using a **trailing coefficient test**. For each selected subset of the set S of irreducible modulo m factors, the trailing coefficient of the product is computed as the product of the trailing coefficients. Only if this product divides the trailing coefficient of A does the algorithm compute the polynomial product and perform a polynomial trial division.

Musser applied his algorithm, implemented on a UNIVAC 1108 computer, and with $v = 5$, to 38 randomly generated polynomials with degrees ranging from 10 to 20 and with coefficient sizes ranging from 2^7 to 2^{21} . Some were irreducible and others were reducible, having been generated as products of polynomials of degrees 2, 3 and 5, or 3, 5 and 7. The computing times for these examples ranged from 0.38 to 27.36 seconds. The irreducible polynomials were often quickly detected using the degree sets; but in several cases Hensel's algorithm had to be applied and in these cases the computing times were much the same as for the reducible polynomials. It is also interesting to observe that in none of the cases did the time to proceed

from the modulo m factorization to the factorization over I take more than 7% of the total time even though this is the part of the algorithm which is least satisfactory from a theoretical viewpoint.

In 1969, Berlekamp devised a new algorithm for the factorization of squarefree univariate polynomials over $\text{GF}(q)$, [2]. Berlekamp's new algorithm has four parts. In part 1, distinct degree factorization is applied. In part 2, the factorization of a squarefree polynomial over $\text{GF}(p)$ whose irreducible factors are all of the same degree is reduced to the factorization of some squarefree polynomials over $\text{GF}(q)$ whose factors are all linear. In part 3 the factorization of a squarefree product of linear factors over $\text{GF}(p^m)$ is reduced to the same problem over $\text{GF}(p)$. Finally, part 4 is an algorithm for factoring a product of distinct linear factors over $\text{GF}(p)$. The computing time for each of the first 3 parts of Berlekamp's new algorithm is dominated by a polynomial function of $n = \deg(A)$ and $L(q)$. The algorithm of part 4 has an average computing time dominated by $n^3 L(p)^3$ but its maximum computing time may be codominant with $n^3 p L(p)^3$. (A slightly different version of the algorithm has a maximum computing time dominated by $n^3 p^{1/4} L(p)^{9/2}$, but its average computing time is not known to be dominated by a polynomial function of n and $L(p)$.)

We have discussed Hensel's algorithm above as a means of obtaining a factorization of A modulo p^k from a factorization of $\phi_p(A)$ over $\text{GF}(p)$, where A is a polynomial over I , p is a prime integer, and $\phi_p(A)$ is squarefree. More generally, Hensel's algorithm is applicable whenever A is a polynomial over a unique factorization domain \mathcal{J} , p is an irreducible element of \mathcal{J} , ϕ_p is the natural homomorphism of \mathcal{J} onto $\mathcal{J}/(p)$, $\mathcal{J}/(p)$ is a field, and $\phi_p(A)$ is squarefree over $\mathcal{J}/(p)$. If we set $\mathcal{J} = \text{GF}(q)(x_1, \dots, x_{r-1})[x_r]$ and choose p as an irreducible polynomial of degree n in x_r over $\text{GF}(q)$, then $\mathcal{J}/(p)$ is $\text{GF}(q^n)(x_1, \dots, x_{r-1})$ and we obtain thereby a Hensel-Berlekamp algorithm for multivariate polynomials over any finite field. If we set $\mathcal{J} = Q(x_1, \dots, x_{r-1})[x_r]$, $r \geq 1$, and choose p as a polynomial $x_r - a$, $a \in I$, then $\mathcal{J}/(p) = Q(x_1, \dots, x_{r-1})$ and we obtain a Hensel-Berlekamp algorithm for multivariate polynomials over I . However, this approach involves rational function coefficients, and Musser [50], has devised a still more general Hensel algorithm which avoids this difficulty. As yet, none of these multivariate factorization algorithms have been implemented and tried, but the supporting theory and the experience to date with univariate factorization indicate their feasibility.

6. Other operations. B. F. Caviness is developing a SAC-1 system for operations in $G[x_1, \dots, x_r]$, where G is the ring of Gaussian integers. Some interesting problems arise in this endeavour when one attempts to obtain optimal algorithms. For example, the classical algorithm for division of rational integers has a computing time codominant with $L(b)L(a/b)$ for division of a by b . However, the obvious algorithm for division of Gaussian integers has a computing time codominant with $L(b)L(a)$, where $L(c_0 + ic_1) = L(c_0^2 + c_1^2) \sim L(c_0^2) + L(c_1^2) \sim L(c_0) + L(c_1)$. Does there exist an

algorithm for Gaussian integer division, using classical algorithms for rational integer arithmetic, whose computing time is dominated by $L(b)L(a/b)$? Caviness has adapted Lehmer's rational integer g.c.d. algorithm to Gaussian integer g.c.d. calculation [9], and is developing a modular algorithm for g.c.d. calculation in $G[x_1, \dots, x_r]$. There is a potential application of this work in performing symbolic calculations with elementary transcendental functions [8], and in computing the complex zeros of a polynomial.

Algorithms for arithmetic operations on rational functions are quite simply obtained in terms of arithmetic operations on polynomials and polynomial g.c.d. calculation. If \mathfrak{U} is any g.c.d. domain with multiplicative ample set \mathcal{A} and \mathcal{J} is the fraction field of \mathfrak{U} , the elements of \mathfrak{U} can be uniquely represented as pairs (a_1, a_2) such that $a_2 \neq 0$, $\gcd(a_1, a_2) = 1$, and $a_2 \in \mathcal{A}$. If $\mathfrak{U} = \mathfrak{U}_0[x_1, \dots, x_r]$, then \mathcal{A} may be defined from an ample set \mathcal{A}_0 for \mathfrak{U}_0 as discussed earlier. If $A(x_1, \dots, x_r) \in \mathfrak{U}[x_1, \dots, x_r]$ is regarded as an element of $\mathfrak{U}_0[x_1, \dots, x_{r-1}][x_r]$, the **leading \mathfrak{U}_0 -coefficient** of A is defined, recursively, as the leading \mathfrak{U}_0 -coefficient of the leading coefficient of A if $r > 1$, and as the leading coefficient of A if $r = 1$. Then $A \in \mathcal{A}$ just in case its leading \mathfrak{U}_0 -coefficient is in \mathcal{A}_0 . If \mathcal{J}_0 is the fraction field of \mathfrak{U}_0 then $\mathcal{J}_0(x_1, \dots, x_r)$, the fraction field of $\mathcal{J}_0[x_1, \dots, x_r]$, is isomorphic with $\mathfrak{U}_0(x_1, \dots, x_r)$, the fraction field of $\mathfrak{U}_0[x_1, \dots, x_r]$, and it is generally more efficient computationally to use the latter. This is the approach which has been used, for example, with $\mathfrak{U}_0 = I$ in the SAC-1 Rational Function System [19], which also provides rational number arithmetic as the special case $r = 0$.

The obvious algorithms for addition and multiplication in a fraction field are susceptible of some improvements, as was observed by P. Henrici in 1956, [34], for rational numbers. If $a_1, a_2 \in \mathfrak{U}$ and $a_2 \neq 0$, let us write a_1/a_2 for the unique pair (\bar{a}_1, \bar{a}_2) such that $\bar{a}_2 \neq 0$, $\bar{a}_2 \in \mathcal{A}$ and $\gcd(\bar{a}_1, \bar{a}_2) = 1$. The obvious algorithm for multiplication in \mathcal{J} applies the formula $(a_1, a_2) \cdot (b_1, b_2) = a_1 b_1 / a_2 b_2$. Henrici's algorithm instead sets $(\bar{a}_1, \bar{b}_2) = a_1 / b_2$, $(\bar{a}_2, \bar{b}_1) = a_2 / b_1$, and then $(a_1, a_2) \cdot (b_1, b_2) = (\bar{a}_1 \bar{b}_1, \bar{a}_2 \bar{b}_2)$. The obvious algorithm performs two multiplications and one reduction (the “/” operation); Henrici's performs two multiplications and two reductions. To see how the Henrici algorithm can nevertheless be faster, consider the very special case in which a_1, a_2, b_1 and b_2 are pairwise relative prime integers, all of the same length d . If d is very large, most of the time for either algorithm will be used by the reductions. For some constant c , the time for each of the two reductions in the Henrici algorithm will be approximately cd^2 , while the time for the one reduction in the obvious algorithm will be approximately $c(2d)^2$. Hence Henrici's algorithm will be about twice as fast in this case. When \mathfrak{U} is a polynomial domain in several variables, the Henrici algorithm will generally be faster by a much larger factor. There is also a Henrici algorithm for addition, and a Henrici-type algorithm for differentiation of rational functions.

The integral of a rational function is not in general a rational function. That is, not every rational function is the derivative of some rational function. However,

if $\mathfrak{A} = \mathcal{J}(x)$, \mathcal{J} any field, and if $(A, B) \in \mathfrak{A}$ then there exist unique polynomials $C, D, E \in \mathcal{J}[x]$ such that $A/B = C' + (D/\hat{B})' + E/\bar{B}$, with $C(0) = 0$, $D = 0$ or $\deg(D) < \deg(\hat{B})$, and $E = 0$ or $\deg(E) < \deg(\bar{B})$, where $\hat{B} = \gcd(B, B')$ and $\bar{B} = B/\hat{B} = \text{gsfd}(B)$. It follows that A/B is "integrable" just in case $E = 0$. In any case C' is called the **polynomial part** of the integral of A/B and D/\hat{B} is called the **rational part**. In the case $\mathcal{J} = \mathcal{Q}$, $\int (E/\bar{B})$ is a sum of logarithms $\sum_{i=1}^j \alpha_i \log(x - \beta_i)$, called the **transcendental part**. We shall refer to E/\bar{B} as the **remainder** of the integral. There is a classical algorithm due to Hermite for computing the polynomial and rational parts and the remainder. The first phase of Hermite's algorithm performs a squarefree partial fraction decomposition, expressing A/B in the form

$$(32) \quad A/B = F + \sum_{j=1}^k \sum_{i=1}^j G_{i,j}/B_j,$$

where $\prod_{j=1}^k B_j^j$ is the squarefree factorization of B , F is a polynomial, and the $G_{i,j}$ are polynomials with $\deg(G_{i,j}) < \deg(B_j)$ or $G_{i,j} = 0$. In the second phase, each sum $\sum_{i=1}^j G_{i,j}/B_j^i$ is integrated by parts, yielding

$$(33) \quad \int \sum_{i=1}^j G_{i,j}/B_j^i = \sum_{i=1}^{j-1} H_{i,j}/B_j^i + \int H_{j,j}/B_j.$$

Then

$$(34) \quad D/\hat{B} = \sum_{j=2}^k \sum_{i=1}^{j-1} H_{i,j}/B_j^i$$

and

$$(35) \quad E/\bar{B} = \sum_{j=1}^k H_{j,j}/B_j.$$

For the case $\mathcal{J} = \mathcal{Q}$, Ellis Horowitz ([36], [37]) developed a modular version of Hermite's algorithm, which was found to be much faster than a version using arithmetic in \mathcal{Q} . Horowitz went on to show how the polynomials C, D and E could be computed directly by solving a system of linear equations, avoiding both partial fraction decomposition and integration by parts, and obtaining thereby a still faster algorithm for rational function integration. Both the modular version of Hermite's algorithm and the new Horowitz algorithm are implemented for $\mathcal{J} = \mathcal{Q}$ in the SAC-1 system [23].

Michael T. McClellan has recently devised modular algorithms for various operations of linear algebra on matrices over $I[x_1, \dots, x_r]$, $r \geq 0$. The modular algorithms include matrix multiplication, determinant calculation, matrix inversion, and solution of a general matrix equation $AX = B$. Of course, classical methods are available, which when implemented using modular algorithms for polynomial multiplication, division, and g.c.d. calculation, will produce quite efficient algorithms. However, McClellan has produced still more efficient algorithms by interjecting the modular methods at the matrix level. For matrix multiplication and determinant calculation, modular and evaluation homomorphisms may be applied in a straight-

forward manner, reducing the problem to matrices over $\text{GF}(p)$. A modular algorithm for matrix inversion is equally trivial if one uses the formula $A^{-1} = \text{adj}(A)/\det(A)$, computing simultaneously the adjoint and determinant of A and then reducing each element $(A^{-1})_{ij}/\det(A)$ to lowest terms with a modular polynomial g.c.d. algorithm. A modular algorithm for solution of the general matrix equation where the ranks of A and B are unknown is much more difficult. However, McClellan has devised an ingenious method for specifying a particular solution from among all solutions of $AX = B$ and for rejecting all homomorphisms which fail to contribute to the determination of that particular solution. McClellan's algorithm also detects cases in which $AX = B$ is inconsistent and in cases where there are multiple solutions a basis for the null space of A is computed, from which all solutions are easily obtained. McClellan's work is described in [47], [48] and [49]; his algorithms have been incorporated in the SAC-1 system [24].

Approximation of the real zeros of a univariate polynomial is ordinarily regarded as a numerical problem rather than as an algebraic problem, but Lee E. Heindel [32] and [33], has demonstrated the efficacy of an algebraic approach which uses infinite precision arithmetic. Heindel has developed an algorithm which, given as inputs any non-zero univariate integral polynomial A and any positive rational number ε , produces as output a sequence I_1, \dots, I_r of disjoint intervals with rational endpoints, each of length less than ε , such that if $\alpha_1 < \alpha_2 < \dots < \alpha_r$ are the distinct real zeros (either simple or multiple) of A , then $\alpha_i \in I_i$. Heindel's algorithm uses Sturm's theorem, which, for any squarefree integral polynomial B , enables us to determine the number of real zeros of B in any left-open, right-closed interval $(a, b]$. Thus Heindel's algorithm begins by setting $B = \text{gsfd}(A)$, which has the same zeros as A and is squarefree. A **negative p.r.s.** (over any ordered integral domain) is a p.r.s. $B_1, B_2, \dots, B_s, B_{s+1} = 0$ satisfying, for some c_i, d_i and Q_i ,

$$(36) \quad c_i B_i = Q_i B_{i+1} + d_i B_{i+2}, \quad c_i d_i < 0$$

for $1 \leq i < s$. A **Sturm sequence** for B is a negative p.r.s. for which $B_1 = B$ and $B_2 = B'$. If we denote by $V(a)$ the number of variations of sign in the sequence $B_1(a), B_2(a), \dots, B_s(a)$, then by Sturm's theorem the number of zeros of B in $(a, b]$ is $V(a) - V(b)$. Heindel's algorithm computes a primitive Sturm sequence, in which each B_i is a primitive integral polynomial. Beginning with a single interval $(-U, +U]$, Heindel's algorithm bisects intervals containing more than one zero and discards intervals containing no zeros, finally arriving at a sequence of isolating intervals, each containing exactly one zero of B . Some of the isolating intervals may be longer than ε , but Sturm's theorem is not needed to refine these since if $(a, b]$ contains at most one zero, then it contains one just in case $B(b) = 0$ or $B(a)B(b) < 0$.

The computing time of Heindel's algorithm is dominated by

$$m^5 L(d)^2 + rm^4 L(dl)^2 + rm^3 L(dl)^3 + rm^2 L(de),$$

where $m = \deg(A)$, r is the number of real zeros of A , $d = \max\{|A|_1, |\text{gsfd}(A)|_1\}$,

$l = \lceil 1/\lambda \rceil$, $e = \lceil 1/\varepsilon \rceil$, and $\lambda = \min_{1 \leq i \leq r} (\alpha_{i+1} - \alpha_i)$ if $r \geq 2$, $\lambda = 1$ otherwise. By considering the discriminant of $B = \text{gsfd}(A)$, it can be shown that, for $r \geq 2$,

$$(37) \quad \lambda \geq d^{-m} (2U)^{-(m^2-m)/4}$$

where U is an upper bound on the zeros of A . Since d is an upper bound, we have

$$(38) \quad \lambda \geq (2d)^{-m^2/4}.$$

It would be interesting to know whether there is a much sharper lower bound for λ as a function of m and d , as one would expect.

Heindel's algorithm has been implemented in the SAC-1 system, [21]. Actually, it appears to be much faster than the theoretical analysis would suggest. For example, when applied to the Chebychev polynomials of various degrees, with $\varepsilon = 10^{-15}$, the observed computing times in seconds on a UNIVAC 1108 computer were quite nearly proportional to the square of the degree, as follows:

degree	10	15	20	25
time	20	46	103	171

An approach similar to Heindel's can be used to approximate the zeros, real and complex, of any univariate Gaussian polynomial A , i.e., any univariate polynomial with Gaussian integer coefficients. As before, we begin by computing $B = \text{gsfd}(A)$. Let $B = B_0 + iB_1$ where B_0 and B_1 have rational integer coefficients, let $\bar{B} = \text{gcd}(B_0, B_1)$, $C_0 = B_0/\bar{B}$, $C_1 = B_1/\bar{B}$ and $C = C_0 + iC_1$. Then $B = \bar{B}C$ and every real zero of B is a zero of \bar{B} , which can be computed by Heindel's algorithm. C has no real zeros and so by applying the Routh-Hurwitz theorem to C we can compute the number of zeros of C above the real axis. Since the zeros of \bar{B} occur in complex conjugate pairs, we can also determine the number of zeros of \bar{B} in the upper half-plane using Sturm's theorem, and hence the number of zeros of B in the upper half-plane.

A **Routh-Hurwitz sequence** for C is any negative p.r.s. $D_1, D_2, \dots, D_s, D_{s+1} = 0$ in which $D_1 = C_0$ and $D_2 = C_1$. If $V(a)$ is the number of variations of sign in the sequence $D_1(a), D_2(a), \dots, D_s(a)$ then, according to the Routh-Hurwitz theorem, the number of zeros of C in the upper half-plane is

$$(39) \quad \frac{1}{2} \{ \deg(C) + V(\infty) - V(-\infty) \}.$$

In a forthcoming Ph.D. thesis [53], James R. Pinkert will show how the Routh-Hurwitz theorem can be combined with rotations, translations and other devices to determine the number of zeros of a Gaussian polynomial in any rectangle with sides parallel to the axes, and from this will specify a complete algorithm for approximating the zeros to desired accuracy.

In this survey we have concerned ourselves primarily with operations on polynomials with integer, Gaussian integer, or rational number coefficients although, as we have seen, this leads naturally to coefficients from a finite field. Although many

challenging problems remain regarding the most efficient algorithms for such polynomials, the accomplishments of the last decade have been sufficient to justify attention in the years ahead to algorithms for operations on algebraic numbers and on polynomials with algebraic number coefficients.

An obvious first step is to consider arithmetic in an algebraic number field $Q(\alpha)$. If we are given the minimal polynomial of α , that is the unique monic irreducible polynomial A with coefficients in Q satisfying $A(\alpha) = 0$, then this problem is theoretically trivial since $Q(\alpha)$ is isomorphic to the residue class ring $Q[x]/(A(x))$, and each residue class may be represented by its unique element of degree less than $n = \deg(A)$. However, any non-zero polynomial B over Q can be expressed uniquely in the form

$$(40) \quad B = b \cdot \bar{B},$$

where b is a rational number and \bar{B} is a primitive integral polynomial with positive leading coefficient, and one may ask whether this representation leads to more efficient algorithms for arithmetic in $Q(\alpha)$.

If α is a real algebraic number then $Q(\alpha)$ is an ordered field and we may also require an algorithm for the order relation in $Q(\alpha)$. Equivalently, we seek an algorithm which decides whether any given element $B(\alpha)$ of $Q(\alpha)$ is positive, negative or zero. Since $\deg(B) < n$ we have $B(\alpha) = 0$ if and only if $B = 0$. Thus, referring to (40), the problem is to decide whether $\bar{B}(\alpha)$ is positive or negative. If A has more than one real zero, the answer may of course depend on which zero of A is denoted by α . One solution is to specify α by an isolating interval I , that is, an interval with rational endpoints containing α but no other zeros of A . Applying Sturm's theorem to I we can decide whether I contains any zeros of B . If so, we can bisect I and obtain a smaller isolating interval for α . Since $B(\alpha) \neq 0$ we eventually obtain an isolating interval $(c, d]$ for α which contains no zeros of B , and then $\text{sign}(B(\alpha)) = \text{sign}(B(d))$.

Sturm's theorem is applicable to polynomials with real algebraic coefficients, and so there is a potential of extending Heindel's algorithm to polynomials over $Q(\alpha)$. This of course leads to questions about optimal algorithms for computing g.c.d.'s of polynomials over $Q(\alpha)$ and for generating Sturm sequences over $Q(\alpha)$. These and some related problems are currently being investigated by Cyrenus M. Rubald [55].

Rudiger Loos [45], has recently extended some of these ideas to obtain interesting algorithms for arithmetic in the field \bar{R} of all real algebraic numbers. Any element α of \bar{R} is represented by a pair (A, I) where A is the minimal polynomial of α and I is a rational isolating interval for α . To add α and β , represented by (A, I) and (B, J) for example, we first compute the resultant C , with respect to y of $A(y)$ and $B(x-y)$, and the interval $K = I + J = \{a + b : a \in I \& b \in J\}$. $\gamma = \alpha + \beta$ is a zero of C and $\gamma \in K$, but C may not be irreducible and K may contain more than one zero of C . However, I and J can be simultaneously refined until $K = I + J$ is an isolating

interval, C can be completely factored, and then the unique irreducible factor of C which has opposite signs at the endpoints of K has γ as a zero.

Research supported by National Science Foundation Grant GJ-30125X, the Wisconsin Alumni Research Foundation and the Stanford Artificial Intelligence Project.

References

1. A programming language for information processing on automatic data processing systems Comm. A. C. M., No. 10, 7 (Oct. 1964) 591-625.
2. E. R. Berlekamp, Factoring polynomials over finite fields, Bell System Tech. J., 46 (1967) 1853-1859.
3. ———, Algebraic Coding Theory, McGraw-Hill, New York, 1968.
4. ———, Factoring polynomials over large finite fields, Math. Comp. No. 111, 24 (July 1970) 713-735.
5. W. S. Brown, On Euclid's algorithm and the computation of polynomial greatest common divisors, J. Assoc. Comput. Mach., No. 4, 18 (Oct. 1971) 478-504.
6. ———, ALTRAN User's Manual (2nd ed.), Bell Laboratories, Murray Hill, N. J., 1972.
7. W. S. Brown and J. F. Traub, On Euclid's algorithm and the theory of subresultants, J. Assoc. Comp. Mach., No. 4, 18 (Oct. 1971) 515-532.
8. B. F. Caviness, On canonical forms and simplification, J. Assoc. Comp. Mach., No. 2, 17 (April 1970) 385-396.
9. ———, A Lehmer-type greatest common divisor algorithm for Gaussian integers, Paper presented at SIAM-SIGNUM 1972 Fall Meeting, Austin, Texas, Oct. 1972.
10. G. E. Collins, A method for overlapping and erasure of lists, Comm. ACM, No. 12, 3 (Dec. 1960) 655-657.
11. ———, PM-A system for polynomial manipulation, Comm. A. C. M., No. 8, 9 (Aug. 1966) 578-589.
12. ———, Polynomial remainder sequences and determinants, this MONTHLY, No. 7, 73 (1966) 708-712.
13. ———, Subresultants and reduced polynomial remainder sequences, J. Assoc. Comp. Mach., No. 1, 14 (Jan. 1967) 128-142.
14. ———, Computing time analysis for some arithmetic and algebraic algorithms, Proc. 1968 Summer Inst. on Symbolic Math. Comp., pp. 195-231. IBM Federal Systems Center, 1968.
15. ———, The computing time of the Euclidean algorithm, Stanford University Comp. Sci. Dept. Report No. CS-331, January 1973, 17 pages.
16. ———, The SAC-1 list processing system, Univ. of Wisconsin Comp. Sci. Dept. Tech. Report No. 129, July 1971, 34 pages.
17. ———, The SAC-1 integer arithmetic system-version III, Univ. of Wisconsin Comp. Sci. Dept. Tech. Report No. 145, July 1972, 63 pages.
18. ———, The SAC-1 polynomial system, Univ. of Wisconsin Comp. Sci. Dept. Tech. Report No. 115, March 1971, 66 pages.
19. ———, The SAC-1 rational function system, Univ. of Wisconsin Comp. Sci. Dept. Tech. Report No. 135, Sept. 1971, 31 pages.
20. ———, The SAC-1 polynomial GCD and resultant system, Univ. of Wisconsin Comp. Sci. Dept. Tech. Report No. 145, Feb. 1972, 93 pages.
21. G. E. Collins and L. E. Heindel, The SAC-1 polynomial real zero system, Univ. of Wisconsin Comp. Sci. Tech. Report No. 93, Aug. 1970, 72 pages.
22. G. E. Collins, L. E. Heindel, E. Horowitz, M. T. McClellan and D. R. Musser, The SAC-1

modular arithmetic system, Univ. of Wisconsin Comp. Center Tech. Report No. 10, June 1969, 50 pages.

23. G. E. Collins and E. Horowitz, The SAC-1 partial fraction decomposition and rational function integration system, Univ. of Wisconsin Comp. Sci. Dept. Tech. Report No. 80, Feb. 1970, 47 pages.

24. G. E. Collins and M. T. McClellan, The SAC-1 polynomial linear algebra system, Univ. of Wisconsin Comp. Sci. Dept. Tech. Report No. 154, April 1972, 107 pages.

25. G. E. Collins and D. R. Musser, Analysis of the Pope-Stein division algorithm, Univ. of Wisconsin Comp. Sci. Dept. Tech. Report No. 55, June 1969, 10 pages.

26. G. E. Collins and D. R. Musser, The SAC-1 polynomial factorization system, Univ. of Wisconsin Comp. Sci. Dept. Tech. Report No. 157, March 1972, 65 pages.

27. J. D. Dixon, A simple estimate for the number of steps in the Euclidean algorithm, this MONTHLY, 78 (1971) 374-376.

28. P. Erdős, On the coefficients of the cyclotomic polynomials, Bull. Amer. Math. Soc. No. 2, 52 (February 1946) 179-184.

29. J. H. Griesmer and R. D. Jenks, SCRATCHPAD/1 — An interactive facility for symbolic mathematics, Proc. Second Symp. on Symbolic and Algebraic Manipulation, pp. 42-58, Assoc. Comp. Mach., 1971.

30. A. D. Hall, Jr., The ALTRAN system for rational function manipulation — a survey, Proc. Second Symp. on Symbolic and Algebraic Manipulation, pp. 153-157, Assoc. Comp. Mach., 1971.

31. A. C. Hearn, Reduce 2: A system and language for algebraic manipulation, Proc. Second Symp. on Symbolic and Algebraic Manipulation, pp. 128-133, Assoc. Comp. Mach., 1971.

32. L. E. Heindel, Algorithms for exact polynomial root calculation, Univ. of Wisconsin Ph.D. Thesis, 1970, 153 pages.

33. ———, Integer arithmetic algorithms for polynomial real zero determination, J. Assoc. Comp. Mach., No. 4, 18 (Oct. 1971) 533-548.

34. P. Henriici, A subroutine for computations with rational numbers, J. Assoc. Comp. Mach., No. 1, 3 (1956) 6-9.

35. K. Hensel, Theorie der algebraischen Zahlen, Chapter 4, Teubner, Leipzig, 1908.

36. E. Horowitz, Algorithms for symbolic integration of rational functions, Univ. of Wisconsin Ph. D. Thesis, 1969, 132 pages.

37. ———, Algorithms for partial fraction decomposition and rational function integration, Proc. Second Symp. on Symbolic and Algebraic Manipulation pp. 441-457, Assoc. Comp. Mach., 1971.

38. S. C. Johnson, Tricks for improving Kronecker's polynomial factoring algorithm, Bell Labs. Report, Murray Hill, N. J., 1966, 22 pages.

39. D. E. Knuth, The Art of Computer Programming, Vol. I: Fundamental Algorithms, Addison Wesley, Reading, Mass., 1968.

40. ———, The Art of Computer Programming, Vol. II: Seminumerical Algorithms, Addison-Wesley, Reading, Mass., 1969.

41. ———, Mathematical analysis of algorithms, Stanford Univ. Comp. Sci. Dept. Tech. Report STAN-CS-71-206, March 1971, 26 pages.

42. L. Kronecker, Grundzüge einer arithmetischen Theorie der algebraischen Grossen, Part I, Section 4, G. Reimer, Berlin, 1882.

43. D. H. Lehmer, Euclid's algorithm for large numbers, this MONTHLY, 45 (1938) 227-233.

44. J. D. Lipson, Chinese remainder and interpolation algorithms, Proc. Second Symp. on Symbolic and Algebraic Manipulations, pp. 372-391, Assoc. Comp. Mach., 1971.

45. R. Loos, A constructive approach to algebraic numbers.

46. J. McCarthy et al., LISP 1.5 Programmer's Manual, M. I. T. Press, Cambridge, Mass., 1962.

47. M. T. McClellan, The exact solution of systems of linear equations with polynomial coefficient

ents, Univ. of Wisconsin Comp. Sci. Dept. Tech. Report No. 136 (Ph. D. Thesis) Sept. 1971, 258 pages.

48. ———, The exact solution of systems of linear equations with polynomial coefficients, Proc. Second Symp. on Symbolic and Algebraic Manipulation, pp. 399–414, Assoc. Comp. Mach., 1971.

49. ———, The exact solution of systems of linear equations with polynomial coefficients, to appear in J. Assoc. Comp. Mach.

50. D. R. Musser, Algorithms for polynomial factorization, Univ. of Wisconsin Comp. Sci. Dept. Tech. Report No. 134 (Ph. D. Thesis) Sept. 1971, 174 pages.

51. A. Newell, J. C. Shaw and H. A. Simon, Empirical explorations of the logic theory machine, Proc. 1957 Western Joint Comp. Conf., pp. 218–230.

52. J. R. Pinkert, SAC-1 Implementation of Toom's Algorithm for Fast Multiplication of Large Integers in Radix Representation, unpublished report, August 1968.

53. ———, Univ. of Wisconsin Comp. Sci. Dept. Ph. D. Thesis, in preparation.

54. D. A. Pope and M. L. Stein, Multiple precision arithmetic, Comm. A. C. M. No. 12, 3 (Dec. 1960) 652–654.

55. C. M. Rubald, Univ. of Wisconsin Comp. Sci. Dept. Ph. D. Thesis, in preparation.

56. J. Xenakis, The PL/1-FORMAC Interpreter, Proc. Second Symp. on Symbolic and Algebraic Manipulation, pp. 105–114, Assoc. Comp. Mach., 1971.

57. H. Zassenhaus, On Hensel factorization, I, J. Number Theory, No. 1, 1 (July 1969) 291–311.

ON THE DISCRETE VERSION OF WIRTINGER'S INEQUALITY

O. SHISHA, Aerospace Research Laboratories, Wright-Patterson AFB, Ohio
(Present address: Naval Research Laboratory, Washington, D.C.)

1. Introduction. Wirtinger's inequality [5, p. 185] states that if $x(t)$ is a real function, absolutely continuous in $[0, 2\pi]$, and satisfying

$$x(0) = x(2\pi), \quad \int_0^{2\pi} x(t) dt = 0,$$

then

$$\int_0^{2\pi} x'^2(t) dt \geq \int_0^{2\pi} x^2(t) dt,$$

equality holding if and only if there are real constants A, B such that, throughout $[0, 2\pi]$, $x(t) = A \cos t + B \sin t$.

It is natural to look for a discrete analog of this result. Such an analog is the following proposition:

THEOREM 1. Let $x_1, x_2, \dots, x_n, x_{n+1} = x_1$ be reals with $\sum_{k=1}^n x_k = 0$. Then

$$(1) \quad \sum_{k=1}^n (x_{k+1} - x_k)^2 \geq 4 \left(\sin^2 \frac{\pi}{n} \right) \sum_{k=1}^n x_k^2.$$

Equality holds if and only if there are real constants A, B such that

$$(2) \quad x_k = A \cos \left[\frac{2\pi}{n}(k-1) \right] + B \sin \left[\frac{2\pi}{n}(k-1) \right], \quad k = 1, 2, \dots, n.$$

Theorem 1 readily implies and is implied by its following complex version:

THEOREM 2. Let $z_1, z_2, \dots, z_n, z_{n+1} = z_1$ be complex numbers with $\sum_{k=1}^n z_k = 0$. Then

$$(3) \quad \sum_{k=1}^n |z_{k+1} - z_k|^2 \geq 4 \left(\sin^2 \frac{\pi}{n} \right) \sum_{k=1}^n |z_k|^2.$$

If $n > 2$, equality holds if and only if (*) there is in the complex plane a regular n -gon w_1, w_2, \dots, w_n such that $z_k = F(w_k)$, $k = 1, 2, \dots, n$, where F is some affine transformation of the plane into itself.

Theorem 2 is due to I. J. Schoenberg [7]. Theorem 1 (along with other, related inequalities) was given, independently, by K. Fan, O. Taussky and J. Todd [3], who indicated also that it extends to the complex case as well as to more general spaces.

Observe that if $n = 1$ or $n = 2$, equality clearly holds in (3). The same is true [7] if $n = 3$, because if w_1, w_2, w_3 is any equilateral triangle in the complex plane, there exists an affine transformation F such that $z_k = F(w_k)$ for $k = 1, 2, 3$. In particular, if n is 1, 2 or 3, equality holds in (1).

The main purpose of the present article is to reprove Theorem 1 using some simple geometric facts which are of interest in themselves.

Inequalities (1) and (3) can be modified, replacing the differences in the left-hand sides by differences of higher order [1, 6]. For the sake of completeness, we give, in Section 3, these generalizations.

2. Some geometry. Our proof of Theorem 1 will be based on the following geometric result, in which $d(a, b)$ denotes the distance between a and b :

THEOREM 3. On a sphere $S: d(c, x) = r$ in an N -dimensional (real) Euclidean space, let P_1, P_2, \dots, P_n ($n \geq 3$) be points such that

$$d(P_1, P_2) = d(P_2, P_3) = \dots = d(P_{n-1}, P_n) = d(P_n, P_1) = d^*,$$

and such that c lies in the convex hull of P_1, P_2, \dots, P_n . Then $d^* \geq a_n$, where a_n is the length of a side of a regular n -gon inscribed in a circle of radius r , namely, $2r \sin(\pi/n)$. If c is an interior point of that convex hull (in the sense of the given N -dimensional space) and $N > 2$, then $d^* > a_n$.

We postpone the proof of Theorem 3 to the end of this section.

Proof of Theorem 1. We may assume that $\sum_{k=1}^n x_k^2 > 0$, and $n \geq 3$. Let $P_1 = (x_1, x_2, \dots, x_n)$, and, for $k = 2, 3, \dots, n$, let

$$P_k = (x_k, x_{k+1}, \dots, x_n, x_1, x_2, \dots, x_{k-1}).$$

In (real) Euclidean n -space E_n consider the sphere

$$S: \|x\| = \left(\sum_{k=1}^n x_k^2 \right)^{\frac{1}{2}}.$$

Then the points P_k lie on S , $\sum_{k=1}^n P_k = (0, 0, \dots, 0)$, and

$$(4) \quad \begin{aligned} d(P_1, P_2) &= d(P_2, P_3) = \dots = d(P_{n-1}, P_n) = d(P_n, P_1) \\ &= \left[\sum_{k=1}^n (x_{k+1} - x_k)^2 \right]^{\frac{1}{2}}. \end{aligned}$$

Therefore, by Theorem 3, (1) holds.

We turn now to study the case of equality in (1). It is a straightforward exercise in analytical geometry that there exist real constants A, B satisfying (2) if and only if P_1, P_2, \dots, P_n is a regular n -gon with center Ω , the origin of E_n . Hence, if there exist such A, B , then $d(P_1, P_2)$ is the length of a side of a regular n -gon inscribed in a circle of radius $(\sum_{k=1}^n x_k^2)^{\frac{1}{2}}$, and therefore (1) holds with equality sign.

Conversely, assume (1) holds with equality sign. Let E be the subspace of E_n spanned by P_1, P_2, \dots, P_n . Then $\dim E$, the dimension of E , is > 1 , for otherwise we would have

$$\sum_{k=1}^n (x_{k+1} - x_k)^2 = d^2(P_1, P_2) = 4\|P_1\|^2 = 4 \sum_{k=1}^n x_k^2 > 4 \left(\sin^2 \frac{\pi}{n} \right) \sum_{k=1}^n x_k^2.$$

Suppose $\dim E$ were > 2 . Note that Ω , as a point of E , is interior to the convex hull of P_1, P_2, \dots, P_n . For, otherwise, we could pass through Ω a hyperplane H (of E) such that all P_k lie in one and the same closed half space determined by H . Since $\sum_{k=1}^n P_k = \Omega$, all P_k would lie on H , whose dimension is smaller than that of E . By applying Theorem 3 (in particular, its last sentence) to the space E and to the sphere $S \cap E$, we would reach (1) with strict inequality. Hence $\dim E = 2$. Observe that for no k can we have $P_k = P_{k+2}$. Indeed, if n is odd, such an equality would lead to $x_1 = x_2 = \dots = x_n$, while if n is even, we would obtain $x_1 = x_3 = \dots = x_{n-1}$, $x_2 = x_4 = \dots = x_n$. Since $\sum_{j=1}^n x_j = 0$, the first possibility contradicts $\sum_{j=1}^n x_j^2 > 0$; the second yields $x_2 = -x_1$, hence $P_1 = P_3 = \dots = P_{n-1} = -P_2 = -P_4 = \dots = -P_n$, which implies $\dim E = 1$. From our assumption of equality in (1) and from (4) it now follows that P_1, P_2, \dots, P_n is a regular n -gon with center Ω . Hence there exist real constants A, B satisfying (2).

For the proof of Theorem 3 we shall need another geometric result:

THEOREM 4. *In an n -dimensional (real) Euclidean space, consider a sphere $S: d(c, x) = r$. Then the length of every closed continuous curve lying on S and containing c in its convex hull is $\geq 2\pi r$. If $n > 2$ and c is an interior point (in the n -dimensional sense) of the convex hull of such a curve, then a strict inequality holds.*

Theorem 4, with $n = 3$, is due to W. Fenchel [4], and the proof we give is very similar to his.

We proceed by induction. The theorem is easily seen to hold for $n = 1$ and $n = 2$. Suppose it holds for some $n (\geq 2)$. In an $(n + 1)$ -dimensional (real) Euclidean space consider a sphere $S: d(c, x) = r$, and a closed continuous curve C lying on S and containing c in its convex hull. Among the finite subsets of C whose convex hull contains c , choose one, F , having a minimal number of points, say k . Then [2, p. 35] since C is connected, $k \leq n + 1$. Thus, there exists a hyperplane H , with $F \subseteq H$, $c \in H$. Now, F lies on the intersection of S with H , a sphere S_1 with center c and radius r in the n -dimensional space H . By a proper replacement of arcs s of C joining pairs of points of F with arcs of great circles, we replace C by a closed continuous curve C_1 , lying on S_1 , whose length L_{C_1} is \leq the length L_C of C . Since $F \subseteq C_1$, therefore c lies in the convex hull of C_1 , and so, by the induction hypothesis, $L_{C_1} \geq 2\pi r$. Therefore $L_C \geq 2\pi r$.

Assume that c lies interior to the convex hull of C . Suppose, first, that $k = 2$; say, F consists of the (distinct) points a_1, a_2 . If we had $L_C = 2\pi r$, then clearly we would have $C = G_1 \cup G_2$ where G_1 and G_2 are semi-great circles joining a_1 to a_2 . But then c cannot be interior to the convex hull of C . Hence $L_C > 2\pi r$. Suppose now $k > 2$. By the minimum property of F , it cannot contain two (distinct) points collinear with c . At least one of the above mentioned arcs s is not a (shortest) arc of a great circle, for otherwise C would lie in H and, consequently, its convex hull would have no interior points. Hence $L_C > L_{C_1} \geq 2\pi r$. This completes the proof.

Proof of Theorem 3. We may clearly assume that $N \geq 2$. The statement $d^* \geq a_n$ is equivalent to the assertion that the (shortest) arc of a great circle on S joining P_1 and P_2 (or any other pair of consecutive P 's, including P_n, P_1) has length $\geq 2\pi r/n$. Multiplying both sides by n , we obtain the equivalent statement that a certain closed curve C joining P_1 to P_2 , P_2 to P_3, \dots, P_{n-1} to P_n , and P_n to P_1 has length $\geq 2\pi r$. But this curve lies on the sphere and contains c in its convex hull; therefore, by Theorem 4, the last inequality holds. Furthermore, if $N > 2$ and c lies interior to the convex hull of P_1, P_2, \dots, P_n , then c is interior to the convex hull of C . Therefore, by Theorem 4, the length of $C > 2\pi r$ which implies $d^* > a_n$.

3. Differences of higher order. Given numbers $x_j, x_{j+1}, \dots, x_{j+p}, p \geq 0$, we set, as usual,

$$\Delta^p x_j = \sum_{k=0}^p (-1)^k \binom{p}{k} x_{j+p-k}.$$

THEOREM 5. Let x_1, x_2, \dots, x_n be reals with $\sum_{k=1}^n x_k = 0$. Define x_k for $k = n + 1, n + 2, \dots$ so that $(x_k)_{k=1}^\infty$ will be of period n . Then for every integer $p (\geq 0)$ we have

$$(5) \quad \sum_{k=1}^n (\Delta^p x_k)^2 \geq 4^p \left(\sin^2 p \frac{\pi}{n} \right) \sum_{k=1}^n x_k^2.$$

If $p > 0$, equality holds if and only if there are real constants A, B satisfying (2)

Proof. True for $p = 0$ (for which (5) always holds with equality sign). Suppose true for some $p \geq 0$. We shall prove it for $p + 1$.

Set $x'_k = \Delta x_k = x_{k+1} - x_k$, $k = 1, 2, \dots, n$. Then

$$\sum_{k=1}^n x'_k = x_{n+1} - x_1 = 0.$$

Extend the definition of x'_k to $k = n + 1, n + 2, \dots$ so that $(x'_k)_{k=1}^\infty$ will be of period n . Then $x'_k = x_{k+1} - x_k$ also for $k = n + 1, n + 2, \dots$. By the induction hypothesis, and by Theorem 1,

$$\begin{aligned} \sum_{k=1}^n (\Delta^{p+1} x_k)^2 &= \sum_{k=1}^n (\Delta^p x'_k)^2 \\ &\geq 4^p \left(\sin^2 \frac{p\pi}{n} \right) \sum_{k=1}^n x_k'^2 = 4^p \left(\sin^2 \frac{p\pi}{n} \right) \sum_{k=1}^n (x_{k+1} - x_k)^2 \\ &\geq 4^{p+1} \left(\sin^2 \frac{(p+1)\pi}{n} \right) \sum_{k=1}^n x_k^2. \end{aligned}$$

Suppose there are real constants A, B such that (2) holds. Then

$$x'_k = A' \cos \left[\frac{2\pi}{n}(k-1) \right] + B' \sin \left[\frac{2\pi}{n}(k-1) \right], \quad k = 1, 2, \dots, n,$$

where A', B' are some real constants. Hence

$$\sum_{k=1}^n (\Delta^p x'_k)^2 = 4^p \left(\sin^2 \frac{p\pi}{n} \right) \sum_{k=1}^n x_k'^2,$$

and by Theorem 1, $\sum_{k=1}^n x_k'^2 = 4(\sin^2(\pi/n)) \sum_{k=1}^n x_k^2$. Therefore

$$\sum_{k=1}^n (\Delta^{p+1} x_k)^2 = 4^{p+1} \left(\sin^2 \frac{(p+1)\pi}{n} \right) \sum_{k=1}^n x_k^2.$$

Conversely, suppose the last equality holds. Then

$$\sum_{k=1}^n (x_{k+1} - x_k)^2 = 4 \left(\sin^2 \frac{\pi}{n} \right) \sum_{k=1}^n x_k^2,$$

and therefore, by Theorem 1, there are real constants A, B satisfying (2).

From Theorem 5 one can deduce its complex analog:

THEOREM 6. Let z_1, z_2, \dots, z_n be complex numbers with $\sum_{k=1}^n z_k = 0$. Define z_k for $k = n + 1, n + 2, \dots$ so that $(z_k)_{k=1}^\infty$ will be of period n . Then for every integer $p (\geq 0)$ we have

$$(6) \quad \sum_{k=1}^n |\Delta^p z_k|^2 \geq 4^p \left(\sin^2 \frac{p\pi}{n} \right) \sum_{k=1}^n |z_k|^2.$$

If $p > 0$ and $n > 2$, equality holds if and only if (*) of Theorem 2 holds.

As in the case of (3), equality always holds in (6) (and hence in (5)), if n is 1, 2 or 3. Of course equality always holds in (6), if $p = 0$.

The author wishes to thank Professors D. Gale and D. J. Newman for their valuable suggestions.

References

1. H. D. Block, Discrete analogues of certain integral inequalities, *Proc. Amer. Math. Soc.*, 8 (1957) 852–859.
2. H. G. Eggleston, *Convexity*, Cambridge Tracts in Mathematics and Mathematical Physics, No. 47, Cambridge University Press, 1958.
3. K. Fan, O. Taussky, and J. Todd, Discrete analogs of inequalities of Wirtinger, *Monatsh. Math.*, 59 (1955) 73–90.
4. W. Fenchel, Über Krümmung und Windung geschlossener Raumkurven, *Math. Ann.*, 101 (1929) 238–252.
5. G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, 2nd edition, Cambridge University Press, 1952.
6. A. M. Pfeffer, On certain discrete inequalities and their continuous analogs, *J. Res. Nat. Bur. Standards Sect. B*, 70B (1966) 221–231.
7. I. J. Schoenberg, The finite Fourier series and elementary geometry, this MONTHLY, 57 (1950) 390–404.

CURRENT TRENDS IN ALGEBRA

GARRETT BIRKHOFF, Harvard University

1. Introduction. Symbolic algebra is much older than many mathematicians suppose; it can be traced back at least to Diophantus of Alexandria (ca. 250 A.D.) and Brahmagupta (ca. 598–665 A.D.). For this early work, see Cajori [7, Arts. 101–5] and Ball [1, pp. 154–6]. Even so-called “modern” algebra is over a century old!

Garrett Birkhoff is the Putnam Professor of Pure and Applied Mathematics at Harvard, where he did his undergraduate and graduate work, was a Junior Fellow in the Society of Fellows, and has served on the faculty since. He has been a Visiting Lecturer at the University of Washington, University of Cincinnati, and the National University of Mexico, and held a Guggenheim Fellowship. He has served as President of SIAM, Vice-President of the AMS, the MAA, and the American Academy of Arts and Sciences, and Chairman of the CBMS. He is a member of the American Philosophical Society and the National Academy of Sciences, and has received honorary degrees from the National University of Mexico, the University of Lille, France, and the Case Institute of Technology.

His extensive publications in modern algebra, fluid mechanics, numerical analysis, and nuclear reactor theory include the books *Hydrodynamics* (Princeton University Press, 1950); *Lattice Theory* (American Mathematical Society Colloquium Publications, 1940, Third Edition 1967); *Survey of Modern Algebra* (with S. Mac Lane, Macmillan, 1941, 1953, 1965); *Jets, Wakes, and Cavities* (with E. H. Zarantonello, Academic Press, 1957); *Ordinary Differential Equations* (Ginn, 1962); *Algebra* (with S. Mac Lane, Macmillan, 1967); *Modern Applied Algebra* (with T. C. Bartee, McGraw-Hill, 1970). *Editor*.

When you realize this, you should not find it too hard to believe that the availability of high-speed computers is giving rise to new trends in algebra. My ultimate aim is to sketch for you, in §§8–10, what I think these trends are. But I wish to lead up to this theme by a brief résumé of the development of algebra as we know it today, over the past several centuries.

2. Classical algebra. The name *modern algebra* was originally intended (in 1930) to signify a contrast with *classical algebra*, which was generally understood to mean the *theory of equations*. This may be defined as the art of solving *numerical problems* by *manipulating symbols*, and seems to have originated with Al-Khwarizmi and other Islamic mathematicians during the period 800–1000 A.D. As we know, its most essential idea consists in replacing each verbal statement about *numerical* quantities by a symbolic *equation*, whose terms can be rearranged and combined by well-established general laws to give a sequence of equivalent but, hopefully, simpler equations. The original equation can be considered as “solved” when the unknown quantity has been isolated on one side of the equality symbol $=$, on the other side of which is some expression involving only known quantities.

Though the word “root” (of an equation) can be traced back to the Sanskrit,^{1,2} and the word “power” (of a number) appears in al-Khwarizmi’s *Algebra* (“al-jabr”), development of classical algebra in its present form was very gradual. The “al-jabr” of the Arabs did not become widely used in Western Europe, and the symbols $+$ and $-$ did not achieve their present significance, until nearly 1500. A major advance followed shortly thereafter, the solution of cubic and quartic equations by radicals being already contained in Cardan’s *Ars Magna* (1545).

For the next two centuries, progress in algebra was mainly³ in connection with its applications: to (analytic) geometry, which gave a vivid meaning to negative numbers, and to the calculus through the use of infinite series. Until after 1750, the significance of imaginary roots and complex numbers remained quite obscure, and even discussions of simultaneous linear equations and determinants were unsystematic and fragmentary.

But from 1750 to 1830, thanks especially to the work of Euler, Lagrange, and Gauss, classical algebra developed rapidly into approximately its present form. Thus the exponential function e^z became defined for all complex z as a power series, and as a result $a^z = e^{z \ln a}$ became well-defined for all positive a and complex z . The “Lagrange resolvent” was also invented by Euler [13, p. 27].

Above all, the Fundamental Theorem of Algebra was recognized as such, clearly stated, and proved. Euler considered its *real* forms, whose equivalence is easily shown. Two of these are:

- (a) Every real polynomial of degree $n > 2$ has proper factors.
- (b) Every real polynomial can be (uniquely) factored into real linear and quadratic factors.

1. All notes are collected together at the end of the paper.

Condition (a) follows for $n < 5$ from Cardan, and for $n = 5$ because every real polynomial of odd degree has real roots. Euler satisfied himself that it was true for all n , but his proof is obscure.

Conditions (a) and (b) are easily shown to be equivalent also to the usual statement of the Fundamental Theorem of Algebra:

(c) Every complex polynomial can be factored into linear factors.

Gauss gave many relatively rigorous proofs of (c) from about 1800 on, and made it clear that all polynomial equations had solutions in terms of complex numbers $x + yi$, $i = \sqrt{-1}$, while the geometrical interpretation of complex numbers as points in the (x, y) -plane gave them a more than symbolic meaning.

Gauss also developed systematic iterative as well as elimination techniques for solving systems of simultaneous linear equations, and the laws of determinants also became generally known—all by 1825 or so.

A few years later, Galois and Abel showed that it was impossible to solve a general equation of the fifth degree by radicals⁴, and after this mathematicians gradually began to turn their attention from the theory of equations to *non-numerical* applications of symbolic algebra (e.g., to groups, vectors and matrices).

MODERN ALGEBRA TO WORLD WAR I

3. “Modern” algebra to 1860. As a result, although real and complex algebra dominated the textbook literature for a full century after 1830, “modern” algebra had already achieved some notable successes by 1860.

Actually, already by 1770, Lagrange was interested in the “symmetric group” of all permutations of n letters and its subgroups, whose relevance to the solution by radicals of the general polynomial equation

$$(1) \quad x^n + a_1x^{n-1} + \cdots + a_n = (x - x_1)(x - x_2) \cdots (x - x_n) = 0,$$

he clearly recognized. A by-product of this interest was the Lagrange Theorem, that the order of any subgroup of a group G divides the order of G .

Ruffini, Galois, and Cauchy made further contributions to the development of group theory before 1845 [13, pp. 45–53]; Galois also made (in 1830) a fundamental contribution to the theory of fields, by constructing a *finite field* of each prime-power order p^r . (For formal definitions of groups and fields, see §3.)

Somewhat earlier Legendre and Gauss (1801) had initiated the study of *commutative rings*, by constructing the ring Z_n of the integers “modulo” n (i.e., in which integral multiples of n are set equal to zero) and the ring $Z[i]$ of all “Gaussian integers” $m + ni$, where $m, n \in Z$ are ordinary integers⁵ and again $i = \sqrt{-1}$. Moreover Gauss, had proved that factorization into primes was *unique* in $Z[i]$.

By 1850, *noncommutative rings* were also being studied. Thus Hamilton introduced his *quaternions* in 1843; since they contain the complex numbers as a special case, they may be called *hypercomplex numbers*. And in the first edition of

his book *Ausdehnungslehre* (1844) H. Grassmann discussed both *vector algebra* (a fairly natural generalization of Descartes' symbolic method for treating geometry) and, somewhat vaguely, hypercomplex numbers in general. These concepts were made much more precise (and their connections with n -dimensional geometry clarified) by Cayley; by Hamilton in the Preface to his book on *Quaternions* (1859); and by Grassmann in the second edition of his book (1878). Moreover, Cayley showed in 1858 [13, p. 84] that the theory of determinants of Vandermonde and Laplace was only one aspect of a much more powerful *matrix algebra*. Matrix algebra is much like ordinary algebra, except that for general matrices A and B , $AB \neq BA$; the multiplication of matrices, like that of quaternions, is *non-commutative*. Indispensable for all pure and applied mathematicians today, matrices were first introduced formally by Cayley in 1858, and gradually revolutionized linear algebra.⁶

Shortly before, two other novel areas of modern algebra had been opened up. In 1854, Boole had published his *Introduction to the Laws of Thought*, in which he showed that a substantial part of Aristotelian logic was described by an analog of ordinary algebra now called "Boolean algebra." This novel "algebra of logic" satisfied not only most of the laws of ordinary algebra, but also the curious identities

$$a^2 = a + a = a \text{ (which today would be written } a \wedge a = a, \vee a = a),$$

$$(a + b)a = a, \text{ and } (a + b)(a + c) = a + bc.$$

4. The axiomatic approach. We have just seen that many of the major branches of so-called "modern algebra" (rechristened "the new math" by the popular press in the post-Sputnik era) were already known to mathematicians by 1860. However, the axiomatic approach to the foundations of algebra did not come until later. Lagrange derived the Lagrange theorem for groups and Galois constructed Galois fields without ever thinking of groups or fields as defined by postulates at all; their assumptions were entirely intuitive! Even the names "commutative" and "distributive" for the corresponding laws of manipulation were not introduced (by Servois) until 1814,⁷ nor the term "associative" (by Hamilton) until 1835.

The emancipation of algebra from exclusive concern with the real and complex fields owes much to the philosophical speculations about algebra of Peacock, Woodhouse,⁸ Hamilton, de Morgan, Boole, and Cayley, but E. T. Bell's claim [3, pp. 180-1] that it was Peacock who: "first perceived common algebra as an abstract hypothetico-deductive science of the Euclidean pattern" goes too far. Though Peacock anticipated Hankel in announcing the "principle of permanence of equivalent forms," his "Symbolical Algebra" is mainly concerned with geometrical applications, and does not even mention axioms or postulates. In these qualities it resembles H. Grassmann's *Ausdehnungslehre* (1844).⁹

The role of axioms emerges much more clearly from the *Formenlehre* of R. Grassmann (1872); the *Operationskreis der Logikkalkul* of E. Schröder (1877); the axiomatic treatments of groups, fields, modules, and ideals by Cayley (1878), Frobenius

and Stickelberger (1879), Dedekind,¹⁰ Weber (1882, 1893), and E. H. Moore; and the independent contemporary work of Benjamin Peirce and his son, C. S. Peirce, at Harvard (1870–1881).¹¹

Influenced by these writings, Peano¹² initiated in 1888 his axiomatic approach to arithmetic, about which I shall say much more later. A decade later, in his *Grundlagen der Geometrie* [9], Hilbert tried to improve on Euclid. He succeeded from the standpoint of rigor, but not from that of pedagogy! Perhaps his most fundamental contribution to axiomatics was his clear formulation of the notions of independence, consistency and completeness for axiom systems.

In 1902, E. H. Moore showed that Hilbert's own axioms were not independent, and during the next ten years E. V. Huntington, L. E. Dickson, and O. Veblen made other painstaking analyses of the independence of postulate systems for groups, fields in general, the real and complex fields in particular, the algebra of logic, and the foundations of geometry. One can get an excellent picture of this work by reading the papers by Moore and Huntington;¹³ for a more colorful if less reliable survey, see [2, Ch. 3].

Partly as a result of such papers, the postulational approach to algebra finally became standard. Mathematicians found that amazingly few and simple postulates, many fewer than those of Euclidean geometry,¹⁴ could provide a sufficient basis for very extensive algebraic theories. For example, all of group theory can be derived from general principles of logic and the following postulates, due to E. V. Huntington (1906).

DEFINITION. A *group* G is a set of elements (to be denoted by small Latin letters), any two of which, say x and y , have a *product* xy which satisfies the following conditions:

G1. Multiplication is *associative*: $x(yz) = (xy)z$ for all $x, y, z \in G$.

G2. For any two elements $a, b \in G$, there exist $x, y \in G$ such that $xa = b$ and $ay = b$.

(We have used Peano's notation $x \in G$ above; it signifies that "the element x is a member of (belongs to) the set G .")

Ingenious arguments can be used to deduce from these postulates various other simple conditions, for example, that: (i) any group G contains a unique "idempotent" element e satisfying $ee = e$, (ii) this element satisfies $ex = xe = x$ for all $x \in G$ (acts as an "identity" for G), (iii) the elements x and y in G2 are uniquely determined by a and b , and so on.

Similarly, the entire theory of fields can be deduced from the following set of postulates, also due to Huntington.

DEFINITION. A *field* is a set F of elements, any two of which have a *sum* $x + y$ and a *product* xy which satisfy the following conditions:

F1. Addition and multiplication are *commutative*:

$$x + y = y + x \text{ and } xy = yx \text{ for all } x, y \in F.$$

F2. Addition and multiplication are *associative*:

$$x + (y + z) = (x + y) + z \text{ and } x(yz) = (xy)z, \text{ all } x, y, z \in F.$$

F3. Multiplication is *distributive* on sums:

$$x(y + z) = xy + xz \text{ for all } x, y, z \in F.$$

F4. For any $a, b \in F$, there exists some $x \in F$ such that $a + x = b$.

F5. If $a + a \neq a$, then there exists some $y \in F$ such that $ay = b$. (Actually, Huntington weakened F5 by adding the condition $b + b \neq b$ to its hypothesis.) (Of course, the hypothesis $a + a \neq a$ is just an indirect way of assuming that $a \neq 0$, necessary here because Huntington wanted to avoid assuming the existence of a "zero" 0 in F .)

The postulational approach to algebra, combined with an awareness of the relevance of all kinds of algebraic systems, stimulated an interest in enumerating *all possible algebraic systems* satisfying specified conditions: all finite fields (Galois had found them all), all groups of given order n , and so on. In this enumeration, one must of course identify all groups (or fields) which are *isomorphic*, that is, whose elements are related by a *bijection which preserves group multiplication* (in fields, which preserves addition *and* multiplication). Such a bijection is called an *isomorphism*.

5. Morphisms and subalgebras. More generally, it is helpful to know when two algebraic systems A and B are related by a (homo)morphism, or mapping $\theta: A \rightarrow B$ which preserves all their defining operations. Finally, it is helpful to recognize the *subalgebras* of A , i.e., the subsets S of A which satisfy all postulates; under these circumstances, A is conversely called an *extension* of S . (Thus the complex field is an extension of the real field.) To test for being a subalgebra, it is usually sufficient to test for *closure* with respect to suitable operations. In a group, for example, a subgroup must contain: (i) the identity, (ii) with any x also x^{-1} ; and (iii) with x and y also xy . In fields, one must require closure under addition, subtraction, multiplication, and division.

The preceding concepts apply to all of the usual kinds of algebraic systems; I shall come back to them in my next lecture.

6. Some deeper developments: 1860–1914. During the same decades that its foundations were being clarified by the postulational method, the scope and depth of algebra grew enormously. I can only indicate very sketchily a few especially remarkable results here.

First, Galois theory became clarified as follows; I shall stick to extensions of the rational field Q to fix ideas, but the results generalize to extensions of any field. Let $F = Q[x_1, \dots, x_n]$ be the field generated by the roots of a polynomial

$$p(x) = (x - x_1)(x - x_2) \cdots (x - x_n) = x^n + a_1x^{n-1} + \cdots + a_n = 0$$

with coefficients $a_k \in Q$. Define the *Galois group* $F(f: Q)$ of F (and of $p(x)$) over Q to be the group of all automorphisms α of F such that $\alpha(x) = x$ for all $x \in Q$. Then the theorem of Galois states that the equation $p(x) = 0$ is solvable by *radicals* in terms of the coefficients (i.e., over Q) if and only if the Galois group $G(F: Q)$ is "solvable" in the following sense.

DEFINITION. Define a *composition series* of a finite group G to be a chain of subgroups of G ,

$$1 < S_1 < S_2 < \cdots < S_r = G,$$

each of which is a maximal normal subgroup of the one following. Form the associated quotient-groups S_k/S_{k-1} . Then G is called *solvable* when these quotient-groups are all *Abelian* (it is equivalent that they all be of prime order).

Second, pure group theory acquired depth. Among the many remarkable theorems about finite groups proved in the half-century 1860–1914, I shall mention only a few. First, it was shown that the set of S_k/S_{k-1} is the same (up to isomorphism and rearrangement) for all composition series (Jordan-Hölder Theorem). Again, it was shown that any group of prime-power order p^n is solvable. Finally, it was shown that if p^n divides the order of a group G , then G has a subgroup of order p^n (Sylow theorem).

Third, in the area of algebraic number theory, Dedekind developed ideal theory, and applied it to generalize the pioneer result of Gauss on the unique factorization of Gaussian integers, to a sweeping unique factorization theorem for any algebraic number field (i.e., any subfield of the complex field C having finite linear dimension over the rational subfield Q). Namely, he showed that factorization into prime ideals is unique.¹⁵

Dedekind's deep interest in ideal theory and in unique factorization into primes also led him to consider the operations of greatest common divisor (g.c.d.) and least common multiple (l.c.m.) from a postulational standpoint. Recognizing their analogy with "and" and "or" in Boolean algebra, he was led to develop and apply the elementary theory of *lattices* ("Dualgruppen"), *modular lattices*, *distributive lattices*, and *vector lattices* in two pioneer papers (1897, 1901), thus founding a major new branch of algebra which contained Boolean algebra as a special case.

7. Linear associative algebras. In 1870, at about the same time that Dedekind was developing ideal theory into a powerful tool, Benjamin Peirce of Harvard made a pioneer study of the systems of "hypercomplex numbers" vaguely adumbrated by Grassmann, Hamilton and Cayley. Peirce began by defining a "linear algebra" over a field F as a set A whose elements are arbitrary linear combinations

$$(2) \quad \mathbf{a} = (a_1, \dots, a_r) = a_1 \mathbf{i}_1 + \cdots + a_r \mathbf{i}_r$$

of r basis elements \mathbf{i}_i , multiplied by some rule of the form

$$(2') \quad \mathbf{a} \cdot \mathbf{b} = (\sum a_l b_m) \mathbf{i}_l \cdot \mathbf{i}_m = \sum a_l b_m \gamma_{lmn} \mathbf{i}_n;$$

the constants γ_{lmn} can be any scalars (elements of F). He called a linear algebra *associative* when the multiplication defined by (2') is associative.

A very notable linear associative algebra is provided by Hamilton's quaternions, which have four basic elements $\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k}$ and hence 64 constants (mostly zero) defined by the rules

$$(3) \quad \mathbf{1} \cdot \mathbf{a} = \mathbf{a} \cdot \mathbf{1} = \mathbf{a} \text{ for all } \mathbf{a}, \quad \mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -\mathbf{1},$$

$$(3') \quad \mathbf{i} \cdot \mathbf{j} = -\mathbf{j} \cdot \mathbf{i} = \mathbf{k}, \quad \mathbf{j} \cdot \mathbf{k} = -\mathbf{k} \cdot \mathbf{j} = \mathbf{i}, \quad \mathbf{k} \cdot \mathbf{i} = -\mathbf{i} \cdot \mathbf{k} = \mathbf{j}.$$

The identities of (3') are clearly those for vector products. The quaternion algebra $R[\mathbf{i}, \mathbf{j}, \mathbf{k}]$ over the real field is also a *division algebra*: any nonzero quaternion $\mathbf{a} = a_0 + a_1 \mathbf{i} + a_2 \mathbf{j} + a_3 \mathbf{k} \neq \mathbf{0}$ has an inverse, given by

$$(3'') \quad \mathbf{a}^{-1} = (a_0 - a_1 \mathbf{i} - a_2 \mathbf{j} - a_3 \mathbf{k}) / (a_0^2 + a_1^2 + a_2^2 + a_3^2).$$

Peirce¹⁶ showed that the complex numbers and the quaternions formed the *only* hypercomplex division algebras over the real field.

Even more important is the *full matrix algebra* $M_n(F)$ of all n^2 -matrices $A = \|a_{lm}\| = \sum a_{lm} e_{lm}$. The basis elements e_{lm} of $M_n(F)$ are multiplied by the rules that

$$(4) \quad e_{lm} e_{l'm'} = \begin{cases} e_{lm'} & \text{if } m = l' \\ 0 & \text{otherwise.} \end{cases}$$

Hence, the constants are given (in a slightly changed notation) by

$$(4') \quad \gamma_{lm, l'm', l''m''} = \begin{cases} 1 & \text{if } l' = m, l'' = l, m'' = m' \\ 0 & \text{otherwise.} \end{cases}$$

From 1870 on, many mathematicians tried to classify linear associative algebras over the real and complex field, using the Fundamental Theorem of Algebra as a tool where convenient. Papers by Frobenius (1878, 1903), Molien (1893), and Cartan (1898) were especially noteworthy.¹⁷

In a remarkable paper published in 1907, Wedderburn showed that most of the structure theorems of Cartan and Frobenius could be proved for linear associative algebras over an arbitrary field! In particular, he proved the following basic results, whose precise meaning will be explained below. For further details, see [3, Ch. 11]. Wedderburn himself stated that "Most of the results contained in the present paper have already been given, chiefly by Cartan and Frobenius, for algebras over the rational field."

(i) Any linear associative algebra is the direct sum (in the vector space sense) of a "semisimple" subalgebra and a unique "nilpotent" invariant subalgebra;

(ii) The semi-simple summand in (i) is the direct sum of "simple" linear associative algebras, in a unique way;

(iii) Each *simple* summand in (ii) is, for some n , the “full matrix algebra” $M_n(D)$ of all $n \times n$ matrices $A = \|a_{ij}\|$ with entries a_{ij} in a suitable “division algebra” D over F , the field of scalars of the original linear associative algebra.

To explain (i), we recall that a linear algebra is called “nilpotent” when, for some finite integer n , all products $a_1 a_2 \cdots a_n$ of n factors vanish. A “subalgebra” of an algebra is a subset closed under addition and multiplication (as well as linear combination over the field of scalars); such a subalgebra K is called “invariant” when $k \in K$ implies $ak \in K$, and $ka \in K$ for any element a , even if not in K ; this is the condition that K be an *ideal* in the sense of ring theory.

Not all linear algebras are associative. The most important family of non-associative algebras is the family of *Lie algebras*. In these, the associative law is replaced by the following three identities:

$$[aa] = 0, [ab] + [ba] = 0, [[ab]c] + [[bc]a] + [[ca]b] = 0,$$

true for all a, b, c . In the 1870's, Lie had shown that real and complex Lie algebras provided the key to the understanding of *continuous groups* based on a finite number of parameters. It was therefore most remarkable that Killing (1888–1890) and Élie Cartan (1894), were able to prove that Lie algebras satisfied structure theorems somewhat analogous to those for associative algebras stated above—and to determine all “simple” Lie algebras over C . This work of Killing and Cartan on the structure of Lie algebras came *before* the analogous work of Molien and Cartan on linear associative algebras^{17a}.

One can, perhaps, summarize the preceding developments in the statement that more was known about “modern algebra” by research algebraists in 1914 than most Ph.D's know today. However, algebra was still regarded as subordinate to classical analysis, and the complex field reigned supreme. Thus, of the two advanced texts on algebra (as distinguished from number theory) most widely used in 1900, Weber's began with a chapter on algebraic functions and Serret's with one on continued fractions!

8. Symbolic logic to Gödel. In retrospect, it seems not too surprising that the dramatic successes of 19th century algebraists and logicians should have encouraged some imaginative mathematicians to develop a *symbolic logic* which would reduce all theorem-proving to mechanical symbol manipulations according to prescribed rules or “axioms.” Actually, this idea goes back at least to Leibniz, who envisioned around 1700 symbolic methods capable of “increasing the power of reason far more than any optical instrument has ever aided the power of vision.” To his fertile mind, the powerful symbolic algebra of the differential and integral calculus (much of which he invented) must have seemed a direct confirmation of the potentialities of symbolic methods.

The symbolic approach was developed tremendously by Peano from 1889 on. His main contributions to it may be found in his *Formulario Matematico* (5th ed.,

1908), whose preface states that: "All progress in mathematics is in response to the introduction of symbols (ideographic signs). ... Among two symbolic systems, the one with fewer symbols is, in general, preferable. But the fundamental use of [symbolic methods] is to facilitate calculation." The preface continues with a review of the origin of various symbols, including $+$, \times , D (derivative), \int , and those for vector and Boolean algebra. It then proposes for general adoption the symbols \in (for membership) and \exists (there exists). Peano claims that with these, and a handful of other symbols and symbolic conventions, all mathematics can be presented in symbolic form.¹⁸

Actually, Peano was not the first mathematician to conceive of a purely symbolic mathematics. In 1634, Hérigone had written in the Preface of his *Cursus Mathematicus*: "I have invented a new method of making demonstrations, brief and intelligible, without the use of any language," and his symbolic style was adhered to by Wallis (1656) and Barrow (1655, 1660).¹⁹

Peano then substantiates his claim by 386 pages of text containing symbolic synopses of: (1) Mathematical Logic, (2) Arithmetic, (3) Algebra, (4) Vectors ("Geometry"), (5) Limits, (6) Derivatives, (7) Integrals. Most successful are Parts (1) and (2); the latter contains Peano's celebrated construction²⁰ of the nonnegative integers by his "successor function": $1 = 0+$, $2 = 1+$, $3 = 2+$, ..., and his derivation of the laws of arithmetic from it is superb. In 70 additional pages, Peano extends his purely symbolic treatment to plane curves, differential equations, and various other topics.

However, Peano's *Formulaire* must be viewed as primarily a thought-provoking *tour de force*, in spite of its wealth of ideas and insights. Nowhere does he list the rules of symbol-manipulation for passing from one formula to the next; he fails to provide a system of axioms for logic. His proofs, like Euclid's in geometry, can only be verified by attributing *meanings* to words.

This major gap was filled by Whitehead and Russell in their three volume masterpiece *Principia Mathematica* [18]. Here they specified carefully the symbol-manipulations ("rules of inference") which can be used infallibly in passing from hypotheses to conclusions in symbolic logic (mathematical reasoning).

Using their specified rules of inference as "axioms" for symbolic logic, Whitehead and Russell showed that one can paraphrase symbolically at least the construction of the real field R from the positive integers, as well as much of set theory and arithmetic. These major achievements were presented as empirical evidence supporting the thesis that *all mathematical theorem-proving can be reduced to mechanical symbol-manipulations* (i.e., to pure symbolic logic).²¹

Nobody disputes the claims of Whitehead and Russell, that their rules of inference for "Peanese" (the symbolic language of Peano), are (i) infallible subject to restrictions stated in English in their text, and (ii) sufficient for much of elementary mathematics. However, the actual mathematical coverage of *Principia* (in nearly

2000 pages!) is far less than Peano's, and it cannot be said that their symbolic methods used "increase the power of reason;" I think they *decrease* it, probably for psychological reasons.²²

9. Hilbert and Gödel, 1918–31. Because of its capability of replacing special axioms for the different branches of mathematics by theorems (cf. *Principia Mathematica*, Preface, first paragraph), Hilbert said in 1918 that "Russell's Axiomatization of Logic is the crowning achievement of axiomatics."²³ And Hilbert spoke with authority, as the man who had rigorized the axioms of Euclid in his famous *Grundlagen der Geometrie* only 20 years before. I quote from the introduction to this book:

"Geometry—like number theory—requires for its deductive (*folgerrichtige*) construction only a few basic theorems (*Grundsätze*). These theorems are called *axioms*,²⁴ and their connected development has had numerous treatments since Euclid The following book is a new attempt to develop the simplest possible complete axiom system for geometry ... so as to clarify the significance of the different groups of axioms and the consequences of the individual axioms."

In much the same spirit of *axiomatic analysis*, Hilbert and his collaborators, especially his co-authors W. Ackermann and P. Bernays,²⁵ made after 1918 major efforts to prove *deductively* (by metamathematical arguments) the adequacy of the axioms of Whitehead and Russell (the evidence of *Principia Mathematica* was empirical). They focused attention on two main questions:

(i) Are these axioms *contradiction-free*, i.e., using them, is it impossible to prove both p and its contradiction $\sim p$?

(ii) Can one test the truth or falsity of any given proposition (e.g., of arithmetic) in a finite number of steps?

Hilbert may have been attracted to these questions partly because he had established an analog of the first in his *Grundlagen der Geometrie*, by using Cartesian geometry as a model for Euclidean plane geometry, and of the second for polynomial ideal bases by general transfinite arguments using the "ascending chain condition."

Question (i) was given a positive answer by Ackermann (who had earlier proved the redundancy of one of the Whitehead-Russell axioms) and von Neumann in 1927, *under suitable restrictions*. These restrictions, which are quite technical,²⁶ seemed quite harmless at first sight, and led to a feeling of optimism about Hilbert's program in the years 1927–1930.

Question (ii), the *Entscheidungsproblem* or Decidability problem, was however given a negative answer, even for arithmetic propositions, by Gödel in 1931. By an ingenious use of metamathematical reasoning, ultimately based on Cantor's diagonal construction, he inferred from this undecidability the *incompleteness* of the Whitehead-Russell-Hilbert system in the following sense. Assuming as true the additional *consistency axiom*, that "false formulas are unprovable," one can prove a number-theoretic formula which would not be provable without it. It is a corollary that one

cannot prove that Hilbert's axioms are contradiction-free, so that in particular Question (i) is undecidable.

Thus Gödel's paper shattered Hilbert's high hopes. To quote Hermann Weyl²⁷: "Gödel enumerated the symbols, formulas, and sequences of formulas in Hilbert's formalism in a certain way, and thus transformed the assertion of consistency into an arithmetic proposition. He could show that this proposition can neither be proved nor disproved within the formalism. This can mean only two things: either the reasoning by which a proof of consistency is given must contain some argument that has no formal counterpart within the system, i.e., we have not succeeded in completely formalizing the procedure of mathematical induction; or hope for a strictly 'finitistic' proof of consistency must be given up altogether. When Gentzen (1936) finally succeeded in proving the consistency of arithmetic he trespassed those limits indeed by claiming as evident a type of reasoning that penetrates into Cantor's 'second class of ordinal numbers'."

Gödel's result ended abruptly a half-century of optimism about symbolic logic, at least as formalized by Peano, Whitehead, and Russell. It showed that their formalizations were incapable of resolving the paradoxes and ambiguities of Cantor's theory of infinite sets.²⁸

THE REIGN OF MODERN ALGEBRA, 1930-1970.

10. The rise of "modern" algebra. Just before Gödel shattered the high hopes of symbolic logicians for formalizing all mathematics in terms of "Peanese", van der Waerden's *Moderne Algebra* (1930-31) precipitated a new revolution. The goal of this brilliantly written book is clearly stated in its preface.

"The 'abstract', 'formal' or 'axiomatic' direction, which has given to algebra renewed momentum,²⁹ has above all led to a series of new concepts in *group* theory, *field* theory, *valuation* theory, and the theory of *hypercomplex numbers*, to insight into new connections and to far-reaching results. The main aim of this book is to introduce the reader into this new world of concepts."

As I have indicated, both the axiomatic approach and much of the content of "modern" algebra dates back to before 1914. However, even in 1929, its concepts and methods were still considered to have marginal interest as compared with those of analysis in most universities, including Harvard. By exhibiting their mathematical and philosophical unity, and by showing their power as developed by Emmy Noether and her other students (most notably E. Artin, R. Brauer, and H. Hasse), van der Waerden made "modern algebra" suddenly seem central in mathematics. It is not too much to say that the freshness and enthusiasm of his exposition electrified the mathematical world—especially mathematicians under 30 like myself.

In particular, it made classical *real and complex algebra* seem passé, or at least a part of analysis and not of "algebra" in the true sense. This view is exemplified in

Moderne Algebra, where the real and complex fields are not even *defined* until after Galois theory has been presented, and the existence and uniqueness of a smallest algebraically closed extension of *any* field (Steinitz, 1910) are proved purely algebraically (by transfinite induction). What a contrast with the texts of Weber, Serret, and Perron!

11. Lattice theory. This new attitude was a major stimulus in the rebirth of lattice theory, which had lain dormant since the pioneer papers of Dedekind. In 1933, I wrote that lattice theory provided "a point of vantage from which to attack combinatorial problems in ... abstract algebra."³⁰ And by 1938, enough progress had been made in applying it to logic, algebra, geometry, probability, measure and integration theory, and functional analysis to cause the American Mathematical Society to hold a symposium on the then very fresh subject.³¹

12. College algebra. The displacement of classical algebra by modern algebra took time. Thus it was not until after World War II that modern algebra became popular at the college level in our country—a popularity due partly to the *Survey of Modern Algebra* which Mac Lane and I had published in 1941. Actually, our approach seems quite conservative today! Thus, unlike van der Waerden, we presented the essentials of the theory of equations before defining groups, and the theory of real and complex matrices (including the Principal Axis Theorem for symmetric and Hermitian matrices) with geometric applications before Galois theory. We also included Boolean algebra, thinking it essential for students to understand the algebra of sets and logic; I shall return to this later.

13. Bourbaki's influence. Abstract mathematics, as reformulated by N. Bourbaki³² in his *Éléments de Mathématique*, was popularized in French universities not long after. This many-volume treatise, mostly written in the decade 1945–55, attempts to develop all of (pure) mathematics systematically from the notions of *set* and *function*: it presents the content of mathematics as concerned with abstractly conceived relational *structures over sets* and mappings (especially *morphisms*) between them; cf. Book 1, Ch. 4.

Algebraic structures are treated in this spirit in Book 2, as defined by sets of *elements* with (internal or external) finitary *operations*. The reader is then led authoritatively and surely through a carefully polished and systematic sequence of definitions, examples, and theorems about groups, rings, fields, and most of the other kinds of systems I have mentioned. Other branches of mathematics are treated in much the same style in later books. The net effect is to make mathematics appear as a *polished monolith, built purely deductively from the notions of set and function*.

14. The flowering of abstract algebra. The enthusiasm generated by van der Waerden's book, reinforced in the ways that I have described, has given rise to an unprecedented flowering of all aspects of abstract algebra over the past 40 years. In

particular, the theories of *groups*, *rings* and *fields* (to which the bulk of *Moderne Algebra* was devoted) have achieved new levels of depth and sophistication, of which perhaps the most dramatic example is the result that *every finite group of odd order is solvable*. This result, proved by Thompson and Feit in over 200 pages of very technical reasoning, had long been conjectured—but to prove it would have seemed hopeless to most mathematicians in 1930.

The last 40 years have also seen the theories of Lie, Jordan, and multilinear algebras mature to a point that makes what was known in 1930 seem amateurish if not naive. The same is true of lattice theory, semigroup and quasigroup theory, category theory, and homological and combinatorial algebra, all of which were either unknown or nearly so in 1930. Finally, algebraic geometry has become rigorized as a new branch of axiomatic algebra, based securely on deep results about commutative rings and their ideals and valuations.³³

15. Wider repercussions. The tidal wave generated by enthusiasm about abstract algebra had wider repercussions. Thus to young men in 1930, like myself, van der Waerden's book made *classical analysis* stemming from the calculus ("*analyse infinitésimale*"), which had dominated mathematics for over two centuries, suddenly seem old and tired. Indeed, the abstract approach adopted by van der Waerden for algebra soon became fashionable in functional analysis and topology. The idea that all mathematics could be viewed as topological algebra gained a strong impetus from the solution of Hilbert's Fifth Problem, which showed that the hypothesis of differentiability could be replaced by mere continuity in the theory of Lie groups: any locally Euclidean continuous group is isomorphic to an analytic Lie group [22, p. 184]. Even research on partial differential equations, the traditional stronghold of the applied mathematician, has increasingly centered around the quest for new abstract concepts permitting one to prove extremely general existence and uniqueness theorems.

Partly because of such shifts in emphasis, by 1960 most younger mathematicians had come to believe that all mathematics should be developed axiomatically from the notions of set and function, and this approach had come to seem no longer modern but classical! By 1959, van der Waerden had changed his title from "*Moderne Algebra*" to "*Algebra*." And in the 1960's, Mac Lane and I wrote another "*Algebra*" which went further in the direction of abstraction, by organizing much of pure algebra around the central concepts of morphism, category, and "universality." The "universal" approach to algebra, which I had initiated in the 1930's and 1940's stressing the role of lattices, was developed much further in two important books by Cohn and Grätzer. In a parallel development, Lawvere (1965) proposed "The category of categories as a foundation for mathematics," beginning with the statement³⁴:

In the mathematical development of recent decades one sees clearly the rise of the conviction that the relevant properties of mathematical objects are those which can be stated in terms of their abstract structure rather than in terms of the elements which the objects were thought to be made of. The question thus naturally arises whether one can give a foundation for mathematics

which expresses wholeheartedly this conviction concerning what mathematics is about, and in particular, in which classes and membership in classes do not play any role. Here by "foundation" we mean a single system of first-order axioms in which all usual mathematical objects can be defined and all their usual properties proved. A foundation of the sort we have in mind would seemingly be much more natural and readily-usable than the classical one when developing such subjects as algebraic topology, functional analysis, model theory of general algebraic systems, etc.

16. The "new mathematics" of 1960. In the post-Sputnik era of the early and middle 1960's, enthusiasm went even further. Especially in the United States, a vogue developed for exposing school children to formal concepts of set, function and axiom often only half-appreciated by their teachers! Its proponents encouraged the spread of the myth that these constituted a "New Mathematics," unknown fifty years earlier. One ostensible aim of this vogue was to indoctrinate young people so that they could fill a supposed shortage of mathematical teachers and research workers. This seemed highly desirable at a time when our postwar "baby bulge" and prosperity was quadrupling of the demand for college teachers of mathematics, while an unquestioning faith in the value of basic science was increasing the support for research in pure mathematics at a rate of 10–15 per cent annually. But as of 1972, it all seems strangely out-of-date!

To summarize, algebra developed harmoniously during the years 1930–60, with its main stream flowing smoothly, swiftly, and finally triumphantly in the channels I have described. Some measure of its triumph may be found in the fact that, whereas three of the first four Fields medals were awarded in Analysis (in 1936 and 1950), three of the four awarded in 1970 were in Algebra.

However, in the last 5–10 years, powerful new currents have become apparent. Some of these have arisen as countercurrents to extremism; thus René Thom has recently written a thought-provoking article entitled '*Modern' Mathematics: An Educational and Philosophical Error*,'³⁵ in which he urges that geometry should replace algebra because "any question in algebra is either trivial or impossible to solve. By contrast, the classic problems of geometry present a wide variety of challenges."

However, I do not wish to dwell on the exaggerations of a decade which most of us recall with nostalgia. Extreme abstraction in research circles, attempts to inculcate premature sophistication in children, and uncritical expansionism in basic physical science have provoked reactions which by now threaten to go too far in the opposite direction.

Instead, I wish to describe four *positive* current trends in algebra which, in my opinion, hold great promise for the future.

FOUR COMPUTER-INFLUENCED CURRENT TRENDS

17. The new numerical algebra. Already in the 1940's a new revolution was brewing, whose ultimate implications for mathematics are unpredictable. Namely,

the construction of efficient *high-speed digital computers* was making it feasible to solve mathematical problems whose effective solution would have previously been prohibitively costly and time-consuming. To many mathematicians, including myself, it had become evident by 1950 that the resulting *revolution in applied mathematics* would open up challenging new areas for basic research. In particular, since digital computers can only represent real numbers to a *finite* number of significant digits, and can only represent values of real functions at a *finite* number of points (approximate "nodal values" at "mesh-points"), their use in solving differential equations (e.g., from physics or engineering) requires a very careful *numerical analysis* of *roundoff* and *truncation* errors.³⁶

Thus, to actually *solve* a system of differential equations (to a desired approximation), one usually first replaces it by an approximating system of *algebraic equations* (obtained perhaps by finite difference or finite element methods), whose unknowns typically represent nodal values at mesh-points, which is then solved (also approximately) on a digital computer. I shall say nothing about this first step of *discretization* here, because the theorems in numerical analysis and approximation theory required to justify it belong to classical analysis and not to algebra. Suffice it to say that it often leads to very large matrices and associated systems of simultaneous linear equations, which may involve 50,000 or more unknowns! The main problem is to solve these efficiently.

These matrices typically have many special properties, which must be exploited to achieve efficiency. They are usually very *sparse* (have mostly zero entries), and often symmetric, or symmetrizable by permutations or linear transformations. Their diagonal elements may be "dominant" (i.e., at least as great as the sum of the absolute values of the other entries), and they may have positive diagonal and negative off-diagonal entries. Matrices having all of the above properties are essentially what are called *Stieltjes matrices*; they arise naturally in *network flow problems*.

One usually wants to either: (i) *solve* the linear system (written symbolically $Ax = b$), or (ii) determine *eigenvalues* of A (the former are of course the roots of $|A - \lambda I| = 0$). As regards (i), most mathematicians imagined in 1940 that large linear systems should be solved (if at all!) by *Gaussian elimination*, and that the rest was sheer drudgery. A few eminent analysts (including Gauss, Jacobi, and von Mises) had appreciated the value of *iterative* methods (also used by Gauss) and had studied their rates of convergence, but these methods were (and still are!) totally ignored in textbooks on "linear algebra." Similar remarks apply to eigenvalue problems, where the experience of most mathematicians was limited to 3×3 (if not to 2×2) matrices $A = \|a_{ij}\|$, whose eigenvalues they might have found using textbook formulas to solve the cubic characteristic equation

$$\lambda^3 - (a_{11} + a_{22} + a_{33})\lambda^2 + \beta\lambda - A = 0,$$

where

$$\beta = a_{22}a_{33} + a_{33}a_{11} + a_{11}a_{22} - a_{23}a_{32} - a_{31}a_{13} - a_{12}a_{21}.$$

In practice, such textbook methods are extremely inaccurate and inefficient for most large matrices³⁷, and they were replaced in the 1950's by new algorithms, whose invention and analysis created a major new area of "classical" algebra: the *new numerical algebra*. Excellent surveys of what is now known about this area are contained in authoritative books by Varga [17], Wilkinson [19], and Young [20]; every forward-looking young algebraist should at least be cognizant of their contents!

18. Sparse matrices. The past five years have also seen substantial improvements (over Gauss) in *elimination* techniques for solving large systems with sparse coefficient-matrices. In particular, these have drawn on graph theory for ideas; see [15] for a cross-section of current work.

There are many other interesting new areas of research in (real and complex) numerical algebra. I shall just mention three of the most important; references to activity in them may be found in many review journals:

- (a) Finding the roots of polynomial equations of degrees up to 100.
- (b) "Unconstrained" minimization of functions of many variables.
- (c) Linear programming and other techniques for finding minima of functions subjected to "constraints" by equations and inequalities.

Actually these "new" areas also originated in the 1940's, if not earlier. Thus by 1947, linear programming was defined, and the "simplex method" of solving its problems invented by George Dantzig; see p. 20 of G. Hadley's *Linear Programming* (Addison-Wesley, 1962). Moreover, its fundamental techniques were made accessible at the college freshman level by Kemeny, Snell, and Thompson 10 years later, in their popular *Introduction to Finite Mathematics* (Prentice-Hall, 1957).

19. Integer arithmetic. In programming languages for computers, a basic distinction is made between *exact* "integer arithmetic" and *approximate* "real arithmetic." I have omitted the problems of "integer programming" and of solving Diophantine equations on computer in the above discussion, because they involve integer and not real and complex numerical algebra. Nevertheless, activity in these fields represents another strong trend in contemporary numerical algebra.

20. Theory of automata. Although many mathematicians think of high-speed computers as simply "number-crunchers" or supersliderules whose primary mathematical role is to carry out elaborate numerical computations, and although "arithmetic units" may be the most highly organized special pieces of computer "hardware," computers are actually much more versatile. Large general purpose computers are designed to be *universal* instruments, capable of expediting all kinds of "mental" tasks. Much as the Industrial Revolution was made possible by machines which could perform all kinds of "physical" tasks more cheaply and efficiently than human beings, the Computer Revolution is aimed at doing the same with mental tasks. This prospect makes the study of computers especially fascinating. From a mathematical standpoint, partly because general purpose computers are *digital* assemblies of a

finite set of components, their study is based on a new, *purely algebraic* concept which I shall now define axiomatically.

DEFINITION. A *finite state machine* (or “automaton”) M consists of a collection A of “input symbols,” a collection S of “states,” and a collection Z of “output symbols,” related by two operations $v: A \times S \rightarrow S$ and $\zeta: S \times Z \rightarrow Z$. The operation v assigns to each “input symbol” $a \in A$ and “prior state” $s \in S$ a “new state” $v(a, s) \in S$; the operation ζ assigns to a and s a “printout” $\zeta(a, s) \in Z$. More concretely, such a finite state machine M can be thought of as evolving from a specified *starting state* s_0 , *recursively* by $s_k = v(s_{k-1}, a_k)$, and as *printing out* $z_k = \zeta(s_{k-1}, a_k)$ for $k = 1, \dots, n$ in succession. In this way, it converts strings of input symbols or *programs* a_1, a_2, \dots, a_n into printouts z_1, z_2, \dots, z_n .

Abstractly, a finite state machine is clearly just a new kind of algebraic system $M = [A, S, Z; v, \zeta]$. If one simplifies M by ignoring Z and ζ (this is called a “forgetful functor” in category theory), the simplified M just describes the *action of a free semigroup* (the set A^* of all possible input “programs”) *on a set* (the set S of states). The resulting theory of state machines without output fits nicely into axiomatic (or “modern”) algebra and, as has recently been shown,³⁹ so-called “universal algebra” can be applied to it.

21. Turing machines. Quite similar to finite state machines, but a little more complicated, are the “Turing machines” invented by the logician Turing in 1936, before high-speed general purpose digital computers existed. Turing proved that they could indeed carry out most processes of mathematical “thought.” Thus they are capable of printing out the binary or decimal expansion of any “definable” (alias “computable”) real number, such as e , π , or the k th zero of the Bessel function $J_n(x)$, and they can “deduce all the provable formulas of the *restricted* Hilbert functional calculus,” giving all true theorems and no false ones.

Some two decades after Turing showed that his machines could, *in principle*, carry out the kind of mechanical theorem-proving dreamed of by Leibniz, Whitehead and Russell, and Hilbert, Hao Wang did this *in practice*. Namely, he wrote a special program which produced “proofs” in minutes for all the 350 theorems in the predicate calculus with equality that were actually stated in Whitehead and Russell’s *Principia Mathematica*!⁴⁰

22. Computational complexity; optimization. A third and very strong trend in algebra, and indeed in mathematics generally, is a concern with *computational complexity* and with *optimization*. In all *applications* of algebra, of course, the efficiency of symbol-manipulation is a prime consideration, but for many years it was taboo to discuss it in research journals devoted to pure mathematics.

This snobbish taboo against discussing efficiency obscured some very important basic facts. Thus, in the area of mathematical logic, the scholarly books by Whitehead and Russell and the Hilbert school did *not* seriously try to improve the

efficiency of formal deductive schemes, whereas Leibniz and Peano were really trying to (and did; especially Leibniz!) develop symbolic techniques for making mathematical reasoning more efficient and, therefore, more powerful. This difference becomes painfully obvious if one compares the number of symbols required by Whitehead and Russell to derive the basic formal properties of sets and relations, with the number of words needed by mathematicians to get equally far. So far, it is only in the area of the propositional calculus of logic itself, and by using a powerful computer, that mechanical theorem-proving has been realized on a substantial scale (by Hao Wang, see §21).

Having finally recognized the importance of efficiency, mathematical logicians have begun to analyze the "computational complexity" of applying general definitions to particular cases. Their analysis has already borne fruit in the development of shorter procedures for multiplying numbers and matrices.

Concern with computational complexity in algebra has as its ultimate goal, of course, the *optimization* of symbolic methods. In turn, the question of optimization has already suggested a number of basic problems whose solution should be a continuing challenge, stimulating coming generations of *pure algebraists*. Two of these are, respectively: (i) the "shortest form" problem of Boolean algebra, and (ii) the "most efficient coding" problem of information theory.

Other fascinating optimization problems, concerning which surprising discoveries have recently been made, are: (iii) what is the least number of operations on digits required to multiply two n -digit integers? (iv) what is the smallest number of arithmetic operations required to multiply together two $n \times n$ matrices? (v) how can one solve n simultaneous linear equations in n unknowns with the fewest additions, subtractions, multiplications and divisions? I regret that I do not have time to discuss these problems here, and must refer you to the literature ([21] and [11, vol. 2, pp. 258–78]).

23. Combinatorial algebra. A fourth current trend in algebra is towards emphasis on *combinatorial* ideas,⁴¹ and especially on those involving *graphs* or *networks*. This trend is surely due to an intuitive recognition of the fact that digital computers and the deductive procedures of mathematics have a structure whose analysis requires combinatorial methods. As Hermann Weyl wrote in 1949: "The network of nerves joining the brain with the sense organs is a subject that by its very nature invites combinatorial investigation. Modern computing machines translate our insight into the combinatorial structure of mathematics into practice by mechanical and electronic devices."⁴²

From burgeoning elementary courses in "Discrete Mathematics" which are intended to *precede* courses in axiomatic algebra,⁴³ probability and statistics, to the ambitious 7-volume treatise [11] on *The Art of Computer Programming* being written by Donald Knuth, the new emphasis is the same: permutations, combinations,

partitions, generating functions, trees, sorting, searching, recurrences, and difference equations, block designs, and so on. Even a casual reading of the books I have cited makes it very clear that the 200 year reign of the Calculus and Analysis has ended — and that they will continue to be displaced in our colleges by courses in Algebra in the broadest sense of discrete mathematics and the *science* (no longer just art!) of symbol manipulation.

In a sense, this trend continues the revolution begun by van der Waerden, but there has been a major change. No longer do axioms and deductive systems, patterned after Euclid's *Elements*, seem so fundamental. Neither do groups or rings, with their subgroups, subrings and morphisms. Their place is taken by various *relational structures* (including partial orderings and “complexes” in the sense of combinatorial topology) which are far less amenable to the general algebraic techniques which played such a central role in the “modern algebra” of 1930–1960.

Instead, the kinds of algebraic structures (as contrasted with “relational” structures) which are most relevant to digital computers and combinatorics are loops, monoids and lattices (or groupoids, semigroups and semilattices), which were largely ignored by most algebraists in 1930–1960. Loosely speaking, much as *groups* are related to *symmetries*, so *loops* are related to *designs* (or “patterns”), *monoids* to *actions* (e.g., of input instructions on the states of an automaton), and *lattices* to *structure*.

In particular, Rota⁴⁴ and his associates have shown that lattice theory provides a point of vantage from which to attack combinatorial problems in general, and not just those of algebra as I had stated in 1933 (see §7). Going even further, N. S. Mendelsohn has very recently applied concepts of universal algebra to generate combinatorial designs and vice-versa [23, pp. 123–32].

One naturally wonders where all these new trends will lead to. I am myself sure of only one thing: that they will *not* make the classical “modern algebra” expounded in van der Waerden obsolete, any more than this made real and complex algebra or the calculus obsolete. As Knuth emphasizes ([11, vol. 1, p. 1]; see also [1]) the word *algorithm* (or “algorism”) which is so central to the mathematics of computation is just a corruption of the name Al-Khwarizmi, the originator of the word “algebra.”

Indeed, the four current trends in algebra which I have been describing were merely *stimulated* by the consideration of digital computers, in much the same way that the calculus and analysis were stimulated by thinking about geometry, mechanics and mathematical physics. They are simply opening up new areas of mathematics for future generations to study, with an ever increasing variety and richness of interrelations and applications, in which old and new ideas will mingle and be reshaped. Within a few decades, new concepts and trends may well emerge from this mingling and reshaping. Certainly, this kind of continuing evolution is the only thing that can keep algebra perennially a fresh and exciting subject!

Notes

¹ See footnote on page 761.

² For this and other facts, I am indebted to Professor David Pingree of Brown University; Thomas Hawkins, Gian-Carlo Rota, Gerald Sachs, and John Tate made other very helpful comments.

³ A notable exception is provided by the binomial theorem, discovered by Pascal in 1653. For readable accounts of the facts summarized in this section, see Rouse Ball [1] and E. T. Bell [3].

⁴ Their expositions were very obscure; see G. Birkhoff, *Isis* 3 (1973), 260–7. That of Galois was clarified by Betti in 1852.

⁵ We here follow the usual custom of letting Z (for the German “*Zahl*” meaning integer) stand for the set $\{0, \pm 1, \pm 2, \dots\}$. Gauss attributed the consideration of integers mod n (“modular numbers”) to Legendre.

⁶ For penetrating historical surveys of linear and non-commutative algebra, see N. Bourbaki, [6, pp. 78–91 and 120–28]. For a readable summary of Cayley’s contribution, see pp. 102–15 of E. T. Bell [2].

⁷ Gergonne’s *Annales* 5 (1814–15), p. 93; for Hamilton’s ideas, see his *Mathematical Papers*, vol. III, Cambridge Univ. Press, 1967. Leibniz and Cramer had very fragmentary ideas about determinants; see [1, p. 375] and D. J. Struik, *A Source Book in Mathematics*, Harvard University Press, 1969, p. 180.

⁸ R. Woodhouse, *Phil. Trans.* 91 (1801), 89–119; G. Peacock, *Reps. Brit. Assn. Adv. Sci.* 3 (1834), 185–32 and *Algebra*, 2 vols., 1845; A. de Morgan, *Trans. Camb. Phil. Soc.*, 7 (1839) 173–87 and 287–300; G. Boole, *Cambridge and Dublin Math. J.*, 3 (1848) 183–98.

⁹ F. Klein, *Entwicklung der Mathematik im 19ten Jahrhundert*, vol. 1, p. 175, characterized this as “almost unreadable.”

¹⁰ In his supplements to Dirichlet’s *Vorlesungen über Zahlentheorie*, 1863, 1871.

¹¹ Benjamin Peirce, *Linear Associative Algebra*, Boston, 1870; see also *Amer. J. Math.*, 3 (1880) 15–57, and 4 (1881) 97–229 (reprinted from *Proc. Am. Acad. Boston*, 1875).

¹² See his *Collected Papers*, vol. 2, Cremonese, Rome, 1958, p. 134. In the *Amer. Math. Society Semicentennial Addresses*, vol. 2, p. 15, Bell attributed the postulational approach to Peano! Peano was also the first to *number* his theorems.

¹³ Volumes 3–6 of the *Transactions* of the (then young) American Mathematical Society (1902–5) contain a dozen articles on postulate systems by the men named above.

¹⁴ Euclid’s *Elements*, which included “axioms” for magnitudes (algebra) as well as “postulates” for geometry, were written in Alexandria, Egypt, around 300 B. C.; see Ball [1].

¹⁵ For a historical discussion of ideal theory and Dedekind’s work on algebraic numbers, see [3, Ch. 10].

¹⁶ *Op. cit.* supra, pp. 216–29. The same result was proved independently by Frobenius, *op. cit. infra*.

¹⁷ G. Frobenius, *Crelle*, 84 (1878) 1–63, and *Berlin Sitzb.* (1903) 504–37 and 634–5. Wedderburn’s *Lectures on Matrices*, *Amer. Math. Soc.*, 1934, contains a complete bibliography to 1933.

^{17a} See Thomas Hawkins, *Archive for History of Exact Sciences*, 8 (1972) 243–87.

¹⁸ A related symbolic style of writing was used by E. H. Moore in his *Introduction to a Form of General Analysis*, New Haven Colloquium, Yale Univ. Press, 1910.

¹⁹ See F. Cajori, “Past struggles between symbolists and rhetoricians. . .”, *Proc. Int. Math. Congress Toronto* (1924), vol. 2, pp. 937–41.

²⁰ First published in 1889 (*Arithmetices principia nova methodo exposita*).

²¹ The fact that this was so had been airily asserted a decade earlier by Russell in his witty *Principles of Mathematics*, of which *Principia Mathematica* was originally intended to be comprised in a second volume!

²² See G. Birkhoff, “Mathematics and Psychology,” *SIAM Review*, 11 (1969) 429–69.

- ²³ *Werke*, vol. 3, p. 153; *Math. Annalen* 78, 405–15.
- ²⁴ Hilbert is here slurring over Euclid's distinction between "axioms" (for magnitudes in general) and "postulates" (for geometrical entities).
- ²⁵ Of the books [10] and *Grundlagen der Mathematik* (2 vols., 1939), respectively.
- ²⁶ See S. C. Kleene, *Introduction to Metamathematics*, Van Nostrand, 1932, pp. 204–5.
- ²⁷ This MONTHLY, 53 (1946) 1–18. Gödel's original paper was published in the *Monats. Math. Phys.*, 38 (1931) 173–98.
- ²⁸ Careful historical reviews of the question touched on here may be found in N. Bourbaki, [6, Ch. 1], and (by P. Bernays) in Hilbert's *Werke*, vol. 3, pp. 196–217; this volume also contains Hilbert's papers on logic.
- ²⁹ In German, "der die Algebra ihren erneuten Aufschwung verdankt."
- ³⁰ *Proc. Camb. Phil. Soc.*, 29 (1933) 441.
- ³¹ *Bull. Amer. Math. Soc.*, 44 (1938) 793–827.
- ³² A pen-name assumed in 1937 by a group of then young French mathematicians, who wished to overthrow the domination of French mathematics by classical analysts. See this MONTHLY, 57 (1950) 221–32 for authentic statement of Bourbaki's opinions, including the view that the axiomatic method is "a *standardization* of mathematical technique," and that the principal mathematical structures are those of a group, of order, and of a topological space.
- ³³ For example, anyone doing serious research on algebraic "geometry" today is expected to consider the two-volume treatise on *Commutative Rings* by O. Zariski and P. Samuel as standard *preliminary* material, but not to know Newton's classification of real cubic curves!
- ³⁴ F. William Lawvere, "The category of categories as a foundation for mathematics," *Proc. Conf. Categorical Algebra*, La Jolla, 1965 (S. Eilenberg *et al.*, eds.), Springer, 1966.
- ³⁵ *American Scientist*, Nov. – Dec., 1971.
- ³⁶ Mathematicians habituated to exclusively deductive reasoning should realize that, in practice, error analysis relies very heavily on empirical evidence as well as on theoretical principles.
- ³⁷ Though not as nearly inefficient as Cramer's Rule, which is still often the only prescription given to students!
- ³⁸ See Marvin Minsky, *Computation: Finite and Infinite Machines*, Prentice-Hall, 1967.
- ³⁹ G. Birkhoff and J. D. Lipson, "Heterogeneous Algebras," *J. Comb. Analysis*, 2 (1969).
- ⁴⁰ H. Wang, *IBM J. Res. Develop.*, 4 (1960) 2–22. For the general question of the computer as a "brain," see the reference of note 22.
- ⁴¹ Wallis, Tschirnhaus, and Leibniz all recognized before 1700 that combinatorics belonged to algebra. See [13, p. 14] and [21, p. 2].
- ⁴² E. F. Beckenbach (editor), *Applied Combinatorial Mathematics*, Wiley, 1964, p. 537.
- ⁴³ As currently recommended by the CUPM Panel on the Impact of Computing on Mathematics Courses. On an intermediate level, see C. L. Liu, *Introduction to Combinatorial Mathematics*, McGraw-Hill, 1968; on a more advanced level, see M. Hall, *Combinatorial Theory*, Ginn, 1967.
- ⁴⁴ "On the foundations of combinatorial theory," *J. für Wahrsch.*, 2 (1966) 340–68; *Combinatorial geometries* (preliminary edition), M.I.T. Press, 1970; and refs. given there.

References

1. W. W. Rouse Ball, *A Short History of Mathematics*, 3rd ed., Macmillan, New York, 1901.
2. Eric T. Bell, *Mathematics: Queen and Servant of Sciences*, McGraw-Hill, New York, 1951.
3. ———, *The Development of Mathematics*, McGraw-Hill, New York, 1940.
4. Garrett Birkhoff and Thomas C. Bartee, *Modern Applied Algebra*, McGraw-Hill, New York, 1970.
5. Garrett Birkhoff and Saunders Mac Lane, *A Survey of Modern Algebra*, Macmillan, New York, 1941.

6. Nicolas Bourbaki, *Éléments d'Histoire des Mathématiques*, Hermann, Paris, 1960.
7. Florian Cajori, *A History of Mathematical Notations*, 2 vols., Open Court, Chicago, 1928–9.
8. George Grätzer, *Universal Algebra*, Van Nostrand, Princeton, N. J., 1968.
9. David Hilbert, *Grundlagen der Geometrie*, 1899; 2nd. ed., 1901. Authorized translation by E. J. Townsend, Open Court, Chicago, 1902, 1910.
10. David Hilbert and W. Ackermann, *Grundzüge der theoretische Logik*, 4th ed., 1949.
11. Donald Knuth, *Algorithms*, 7 projected volumes, Addison-Wesley, Reading, Mass., 1969.
12. S. Mac Lane and G. Birkhoff, *Algebra*, Macmillan, New York, 1967.
13. Uta Merzbach, "... Development of Modern Algebraic Structures from Leibniz to Dedekind," Ph. D. Thesis, Harvard, 1965.
14. Giuseppe Peano, *Formulario Matematico*, 4th ed., Torino, 1908.
15. Donald Rose and Ralph Willoughby (eds.), *Sparse Matrices and their Applications*, Plenum Press, New York, 1971.
16. B. L. van der Waerden, *Moderne Algebra*, 2 vols., Springer, New York, 1930–31.
17. Richard S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N.J., 1962.
18. Alfred N. Whitehead and Bertrand Russell, *Principia Mathematica*, 3 vols., Cambridge Univ. Press, 1911.
19. James Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1966.
20. David M. Young, *Iterative Solution of Large Linear Systems*, Academic Press, New York, 1971.
21. Garrett Birkhoff and Marshall Hall (eds.), *Computers in Algebra and Number Theory*, SIAM-AMS Proceedings, vol. IV, Amer. Math. Society, 1971.
22. Deane Montgomery and Leo Zippin, *Topological Transformation Groups*, Wiley-Interscience, New York, 1955.
23. W. Tutte (ed.), *Recent Progress in Combinatorics*, Academic Press, New York, 1969.

MATHEMATICAL NOTES

EDITED BY ROBERT GILMER

Material for this Department should be sent to David Roselle, Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803.

EXISTENCE OF FOUR CONCURRENT NORMALS TO A SMOOTH CLOSED HYPERSURFACE OF E^n

BERND WEGNER, Technische Universität Berlin

A four-normals-point of a smooth closed hypersurface of E^n (n -dimensional euclidean space) is a point where at least four normals of the hypersurface intersect. There are several examples of four-normals-points of a closed curve in the plane (compare Chakerian and Stein [2], Deo and Klamkin [3], Guggenheimer [4]). The purpose of this note is to give many examples of four-normals-points of a smooth closed hypersurface F of E^n . We shall prove that every neighborhood of a focal point of F contains a four-normals-point of F if F is an immersed $(n-1)$ -sphere.

The proof uses elementary Morse theory which can be found in Milnor [5] or Cairns and Morse [1].

First there are given some basic results and definitions from elementary Morse theory. Let F be a closed C^2 hypersurface of E^n which is an immersed $(n-1)$ -sphere. Let L_p denote the distance function with center $p \in E^n$, i.e., $L_p(x) = \|x - p\|^2$. The point $x \in F$ is called **critical point of L_p** if the differential $(L_p)_x^*$ of L_p at x vanishes; the critical point x of L_p is called **nondegenerate** (resp. **degenerate**) if the matrix

$$\left(\frac{\partial^2 L_p}{\partial u_i \partial u_j} \Big|_x \right)$$

is nonsingular (resp. singular) where u_1, \dots, u_{n-1} are local coordinates of F around x ; the **index of a nondegenerate critical point x of L_p** is defined to be the maximal dimension of a subspace of the tangent plane of F at x where the bilinear functional represented by

$$\left(\frac{\partial^2 L_p}{\partial u_i \partial u_j} \Big|_x \right)$$

is negative definite. A necessary and sufficient condition for x to be a critical point of L_p is that the normal line of F through x meets p . A point $q \in E^n$ is called **focal point of F with base $x \in F$ and multiplicity μ** if $q = x + v$ is the image of the point (x, v) of the normal bundle of F under the end point map E and the differential $E_{(x,v)}^*$ of E at (x, v) has **rank** $n - \mu < n$ where E is defined by $E(x, v) = x + v$.

Useful characterizations for q to be a focal point of F with base x are the following two (see Milnor [5], p. 32 ff): (1) x is a degenerate critical point of L_q , (2) $\|q - x\|$ is a principal radius of curvature of F in the direction $(q - x)/\|q - x\|$. Thus the set of focal points of a plane curve is exactly the evolute of the curve.

The second characterization also implies that on the normal line of F through x there is only a finite number of focal points with base x (at most $n-1$). The Morse index theorem for the distance function (see [5], p. 37) states that the index of $x \in F$ as a nondegenerate critical point of L_p is exactly the number of focal points of F with base x situated on the segment from x to p , each being counted with its multiplicity.

THEOREM. *If q is a focal point of F , then every neighborhood of q contains a four-normals-point of F .*

Proof. Let q be a focal point of F with base x . Let N denote the unit normal of F in x such that $q = x + rN$ for some $r > 0$. Furthermore let m (resp. M) be the square root of the absolute minimum (resp. maximum) of L_q on F . If $m = r = M$, then F is a sphere and the statement is trivial. In the case $m < r$ take

$$q' = x + (r - \varepsilon)N$$

where $\frac{1}{2}(r - m) > \varepsilon > 0$ such that q' is not a focal point of F with base x . Using

the triangle inequality we see that $L_{q'}(x)$ is not an absolute minimum of $L_{q'}$. Applying the Morse index theorem for the distance function, we conclude that the index of x as a nondegenerate critical point of $L_{q'}$ is less than $n-1$ because q is not contained in the segment from x to q' .

As we shall see below, the point q' would suffice for a four-normals-point if $L_{q'}$ were nondegenerate (i.e., all critical points of $L_{q'}$ are nondegenerate). This is not necessarily the case. Therefore, we have to look for a nondegenerate $L_{q''}$ with q'' in a sufficiently small neighborhood of q' such that the nice properties of $L_{q'}$ are valid for $L_{q''}$. We shall get the existence of such nondegenerate functions as a consequence of the following statement: There exists a real number $\eta > 0$ such that for any point q'' in the η -neighborhood of q' , $L_{q''}$ has a critical point where the value of $L_{q''}$ is neither an absolute maximum nor an absolute minimum. To see the last statement, we regard a connected open neighborhood U of $(x, q' - x)$ in the normal bundle of F such that the restriction $E|U$ of the end point map to U is a diffeomorphism onto an open subset of E^n . The existence of U follows from the Inverse Function Theorem because q' is not a focal point of F with base x and therefore $E^*_{(x, q' - x)}$ has rank n . Thus for every pair $(z, v) \in U$ we have that z is a nondegenerate critical point of L_{z+v} . Choosing the η -neighborhood of q' sufficiently small within $E(U)$ such that the values of $L_{q'}$ at x and L_{z+v} at z do not differ too much for $(z, v) \in U$ and $z + v$ in this neighborhood we get the statement above by a simple geometric argument.

Now we continue the main proof. The set of focal points of F is nowhere dense in E^n (compare Cairns and Morse [1], Theorem 15.3). Thus, there exists a nondegenerate L_p with p as near to q' as we please (and hence to q), having a critical point where the value of L_p is neither an absolute maximum nor an absolute minimum. Thus by the compactness of F L_p has at least three critical points. If C_i denotes the number of critical points of index i of L_p , we have the equality $\sum_{i=0}^{n-1} (-1)^i C_i = \chi$ (see [5], p. 28) where χ is the Euler characteristic of F which is even in our special case. This implies $\sum_{i=0}^{n-1} (-1)^i C_i \equiv 0 \pmod{2}$, and therefore there must be a fourth critical point of L_p . Hence p must be a four-normals-point of F . In the case $r < M$ replace the inequality given above by $\frac{1}{2}(r - M) < \varepsilon < 0$ and minimum by maximum.

COROLLARY. *For a closed C^2 hypersurface H of E^n (which may have self-intersections) which has an even Euler characteristic there exists at least one four-normals-point.*

Proof. If H is an immersed $(n-1)$ -sphere, then this corollary is a direct consequence of the theorem above because H must have a focal point. If H is not an immersed $(n-1)$ -sphere, then every nondegenerate L_p must have more than two critical points (compare Milnor [5], p. 25) and therefore at least four because χ is even. This gives the proof of the corollary because L_p is nondegenerate for almost

all $p \in E^n$. The last case is also contained in the normal arc theorem of Cairns and Morse ([1], p. 281).

We conclude with two remarks. First, we note that the proofs above do not depend on the codimension one and therefore, the same statements are valid for any other codimension.

We may also state that the $(n-1)$ -sphere is the only hypersurface H of E^n with even Euler characteristic which has only one four-normals-point. This remark is seen as follows: By the proof of the corollary, H must be an immersed $(n-1)$ -sphere. Then the theorem above implies that H has only one focal point, i.e., H is the rigid $(n-1)$ -sphere in E^n .

References

1. S. S. Cairns and M. Morse, *Critical Point Theory in Global Analysis and Differential Topology*, Academic Press, New York, 1970.
2. G. D. Chakerian and S. K. Stein, On the centroid of a homogeneous wire, *Mich. Math. J.*, 11 (1964) 189–192.
3. N. Deo and M. S. Klamkin, Existence of four concurrent normals to a smooth closed curve, *this MONTHLY*, 77 (1970) 1083–1084.
4. H. H. Guggenheimer, Geometrical applications of integral calculus, p. 84 (contained in K. O. May, *Lectures on Calculus*, Holden-Day, San Francisco, 1967).
5. J. Milnor, *Morse Theory*, Princeton University Press, Princeton, 1963.

ON A PROBLEM OF BESICOVITCH

B. FISHER, Leicester

In 1917, S. Kakeya [2] raised the following problem:

A line segment AB lying in a plane is to be moved in the plane so that it returns to its original position but with its direction reversed. How should this be done in order that the area swept out by the segment during motion may be a minimum?

In 1928 A. S. Besicovitch [1] gave an answer to this problem, showing that the area swept out could be made arbitrarily small. He showed that this result was a consequence of the following theorem which he had established [3]:

THEOREM. *Let ABC be any triangle. Divide the base AB into 2^n equal parts and join the points of division to the vertex C , dividing the triangle into 2^n elementary triangles. With suitable choice of n it is then possible to translate the elementary triangles along AB in such a way that the area covered by them in their new positions is arbitrarily small.*

The original proof of this theorem by Besicovitch was rather difficult and I give here a more elementary proof.

Proof of the theorem: Suppose the area of triangle ABC is S and the length of AB is c . Divide AB into 2^n equal parts and label the points of division Y_i for $i = 0, 1, 2, \dots, 2^n$, the particular points given by $i = 0$ and 2^n being A and B respectively. Translate each triangle $Y_{2i}Y_{2i+1}C$ for $i = 0, 1, 2, \dots, 2^{n-1} - 1$ a distance $cx/2^n$ along AB , where $0 < x < 1$, to form the triangle $X_{2i}X_{2i+1}C'$ in its new position. The triangle $X_{2i}X_{2i+1}C'$ overlaps the triangle $Y_{2i+1}Y_{2i+2}C$, the sides $X_{2i}C'$ and $Y_{2i+2}C$ intersecting in a point Z_i , to form a triangle $X_{2i}Y_{2i+2}Z_i$ together with two other triangles outside triangle $X_{2i}Y_{2i+2}Z_i$.

It follows by elementary geometry that the area covered by the two overlapping triangles $X_{2i}X_{2i+1}C'$ and $Y_{2i+1}Y_{2i+2}C$ is now

$$\frac{S}{2^n} [1 + (1-x)^2 + 2x^2],$$

the area of triangle $X_{2i}Y_{2i+2}Z_i$ being

$$\frac{S}{2^{n-1}} (1 - x - \frac{1}{2}x^2) < \frac{S}{2^{n-1}} (1 - x)$$

and the area outside this triangle being $Sx^2/2^{n-2}$. Since $Y_{2i}Z_{i-1}$ is parallel to $X_{2i}Z_i$ for $i = 1, 2, \dots, 2^{n-1} - 1$ it follows that the 2^{n-1} triangles $X_{2i}Y_{2i+2}Z_i$ can be translated along AB to form a triangle $A_1B_1C_1$ with each point Z_i moving to the vertex C_1 . The area S_1 of triangle $A_1B_1C_1$ is less than $S(1-x)$ and the area covered by the original 2^n elementary triangles outside triangle $A_1B_1C_1$ is less than $2Sx^2$.

The above process is now repeated on the triangle $A_1B_1C_1$ which is made up of 2^{n-1} elementary triangles, to form a triangle $A_2B_2C_2$ with area S_2 . We have

$$S_2 < S_1(1-x) < S(1-x)^2$$

and the area covered by the original 2^n elementary triangles outside $A_2B_2C_2$ is less than

$$2Sx^2 + 2S_1x^2 < 2Sx^2 + 2Sx^2(1-x).$$

Further repetitions of this process give us a triangle $A_rB_rC_r$ with area S_r . We have

$$S_r < S_{r-1}(1-x) < S(1-x)^r$$

and the area covered by the original 2^n elementary triangles outside triangle $A_rB_rC_r$ is less than

$$2Sx^2[1 + (1-x) + \dots + (1-x)^{r-1}] < 2Sx.$$

Now, for arbitrary $\varepsilon > 0$, put $x = \varepsilon/2S$. The area covered by the original 2^n elementary triangles outside triangle $A_rB_rC_r$ is then less than ε . Further the area of triangle $A_rB_rC_r$ is less than $S(1 - \varepsilon/2S)^r$, which is less than ε for r greater than

some N . Thus if $n > r > N$ the area covered by the 2^n elementary triangles is now less than 2ε , completing the proof of the theorem.

To see how *Takeya's* problem now follows from this theorem the reader is recommended to read *Besicovitch's* original paper.

References

1. A. S. Besicovitch, On *Takeya's* problem and a similar one, *Math. Zeit.*, 27 (1928) 312–320.
2. S. *Takeya*, *Tôhoku Science Reports*, 6 (1917) 71–78.
3. A. S. Besicovitch, Sur deux questions de l'intégrabilité, *Journal de la Société des Math. et de Phys. à l'Univ. à Perm*, 2 (1920) 105–123.

TOPOLOGICAL PROPERTIES OF THE ROW ECHELON FORM

G. P. BARKER, University of Missouri at Kansas City

Motivated by [1] we investigate the topological nature of the set of matrices in row echelon form. We can define an equivalence relation on this set which involves only the elements of the matrices. This equivalence relation seems rather natural, and it is pleasing to find that it has topological significance.

Throughout this note F denotes either the real or the complex numbers, and M denotes the set of $m \times n$ matrices with elements from F . The usual topology on M can be generated by the norm $\mu(A) = \max |a_{ij}|$ for $A \in M$. Next we recall two definitions [2, p. 44].

Let $A \in M$. The **leading entry** of a row of A is the first nonzero entry of that row. Denote by $l(i; A)$ the index of the column of A which contains the leading entry of row i if that row is nonzero. Otherwise, set $l(i; A) = n + 1$. The matrix A is in **row echelon form** if

- (1) the nonzero rows are at the top of the matrix;
- (2) $l(1; A) < l(2; A) < \dots < l(r; A)$, where $r = \text{rank } A$;
- (3) all leading entries are 1;
- (4) $j = l(i; A)$ implies $a_{ij} = 1$ and $a_{kj} = 0$ for $k \neq i$.

Let R denote the set of $m \times n$ matrices in row echelon form with elements from F . The topology on R is the subspace topology induced by M . If $l(i; A)$ is the column index for the leading entry of row i of A and $l(i; B)$ is the same for B , then A and B are **pattern equivalent** if $l(i; A) = l(i; B)$ for $i = 1, 2, \dots, m$. This relation is denoted by $A \sim B$, and it is easily seen to be an equivalence relation.

The norm μ restricted to R determines an open base of the neighborhoods of any point $A \in R$. If A and B are elements of R such that $\mu(A - B) < \varepsilon$ for $\varepsilon < \frac{1}{4}$, say, then A and B are pattern equivalent. In fact, if

$$S(A) = \{B \in R \mid \mu(A - B) < \varepsilon\},$$

where ε is small, say $\varepsilon < \frac{1}{4}$, and if $0 \leq \alpha \leq 1$, then for any C_1 and C_2 in $S(A)$ we have

$$\alpha C_1 + (1 - \alpha) C_2 \in S(A).$$

We can collect these observations as a formal result.

PROPOSITION. *R is locally convex.*

Consequently, R is locally connected so that the components of R are both open and closed (see [3], p. 72).

In analogy with [1] we shall write $A \approx B$ if A and B lie in the same arcwise connected component of R , and $A \sim B$ if A and B lie in the same component of R . Both \approx and \sim are equivalence relations. The equivalence class of A with respect to \approx , \sim , and \sim will be denoted by $P(A)$, $A(A)$, and $C(A)$ respectively.

THEOREM. *The three relations \approx , \sim , and \sim are equivalent; that is, for each $A \in R$ we have $P(A) = C(A) = A(A)$.*

Proof. We first show that $P(A) = A(A)$.

If $B \in P(A)$, then $f(\alpha) = (1 - \alpha)A + \alpha B$ is an arc from A to B which lies entirely within $P(A)$. Hence $P(A) \subset A(A)$. Conversely suppose

$$f(\tau) = \begin{bmatrix} f_{11}(\tau) & \cdots & f_{1n}(\tau) \\ f_{m1}(\tau) & \cdots & f_{mn}(\tau) \end{bmatrix}$$

is a continuous function from $[0, 1]$ to $A(A)$ with $f(0) = A$ and $f(1) = B$. If $B \notin P(A)$, then for some i necessarily $l(i; B) \neq l(i; A)$. We may assume $l(i; B) < l(i; A)$ so that $f_{ij}(0) = 0$ for $j = 1, 2, \dots, l(i; A) - 1$. Each $f_{ij}(\tau)$ is a continuous function and for $l = l(i; B)$ we have $f_{il}(0) = 0$ and $f_{il}(1) = 1$. Consequently we can find a τ_0 with $0 < \tau_0 < 1$ and a k satisfying $l(i; B) \leq k < l(i; A)$ such that for all sufficiently small $\varepsilon > 0$ we have $f_{ik}(\tau_0 + \varepsilon) \neq 0$ while $f_{ij}(\tau) = 0$ for all $\tau \leq \tau_0$ and all $j = 1, 2, \dots, l(i; A) - 1$. Note that $\tau_0 = 1$ is not possible since that would mean that the continuous function $f_{ik}(\tau)$ would map the interval $[0, \tau_0]$ onto the set $\{0, 1\}$. But now we see that for $\varepsilon > 0$ sufficiently small $f_{ik}(\tau_0 + \varepsilon) \neq 1$ so that $f_{ip}(\tau_0 + \varepsilon) = 1$ for some $p < k$ as $\varepsilon \downarrow 0$. This contradicts the fact that $f_{ip}(\tau_0) = 0$. Hence $B \in P(A)$.

We finish the proof by showing that the arc components are both open and closed. If $A_q \in A(A)$ and $B = \lim A_q$, then each A_q is pattern equivalent to A . However, the convergence is coordinatewise so that B must be pattern equivalent to A . Thus $B \in A(A)$, and $A(A)$ is closed. On the other hand if $B \in A(A)$, and as before, if

$$S(B) = \{C \in R \mid \mu(B - C) < \varepsilon\},$$

where $\varepsilon < \frac{1}{4}$, say, then $S(B)$ is a neighborhood of B which is open in R . It is also clear that $S(B) \subset P(A) = A(A)$. Hence $A(A)$ is also open, and so $A(A) = C(A)$.

Acknowledgement. The author would like to thank the referee for several helpful suggestions.

References

1. H. Schneider, Topological aspects of Sylvester's theorem on the inertia of Hermitian matrices, this MONTHLY, 73 (1966) 817-821.
2. H. Schneider and G. P. Barker, Matrices and Linear Algebra, Holt, Rinehart and Winston, New York, 1968.
3. A. Wilansky, Topology for Analysis, Ginn, Waltham, Mass., 1970.

TWO-DIMENSIONAL COMPLETE MONOTONICITY WITH DIAGONALIZATION

C. H. KIMBERLING, University of Evansville

1. Introduction. A continuous function from $[0, \infty)$ into $[0, \infty)$ is **completely monotone** if its derivatives alternate in sign: $(-1)^n f^{(n)}(x) \geq 0$ for $n = 0, 1, 2, \dots$ and all x in $(0, \infty)$. Correspondingly, a sequence $\mu_0, \mu_1, \mu_2, \dots$ of nonnegative real numbers is a **completely monotone sequence** if for each n , all the differences

$$\begin{aligned}\Delta^1 \mu_n &= \mu_n - \mu_{n+1} \\ \Delta^2 \mu_n &= \mu_n - 2\mu_{n+1} + \mu_{n+2} \\ &\vdots \\ \Delta^k \mu_n &= \mu_n - \binom{k}{1} \mu_{n+1} + \binom{k}{2} \mu_{n+2} - \dots + (-1)^k \mu_{n+k}\end{aligned}$$

are non-negative. Our symbols $\Delta^k \mu_n$ just defined follow [2], but not [3].

We shall consider two types of infinite matrices and associated two-place functions. In the first type of matrix, each row and each column forms a completely monotone sequence. The sequence of diagonal elements of such a matrix need not be completely monotone, but additional monotonicity conditions ensure a completely monotone diagonal. Analogous conditions on a two-place function $f(x, y)$ imply complete monotonicity of the one-place function $f(x, x)$.

The second type of matrix arises from the derivatives of a given completely monotone function f on $[0, \infty)$ as follows: the n th row is $(-1)^n f^{(n)}(k)$, $k = 0, 1, 2, \dots$. The diagonal $(-1)^n f^{(n)}(n)$ is then a completely monotone sequence which extends to a completely monotone function.

2. Two-dimensional complete monotonicity. We call (μ_{ij}) a **completely monotone matrix** if

$$(1) \quad \Delta_1^n \Delta_2^m \mu_{ij} \geq 0; \quad n, m = 0, 1, 2, \dots; \quad i, j = 0, 1, 2, \dots$$

For example, if (μ_i) and (v_j) are completely monotone sequences, then $(\mu_i v_j)$ is a completely monotone matrix. Also (μ_{i+j}) is a completely monotone matrix.

We call (μ_{ij}) a **placewise completely monotone matrix** if each of its rows and columns is a completely monotone sequence. Clearly a completely monotone matrix

is a placewise completely monotone matrix. The converse fails, for example, when $\mu_{ij} = (i+1)^{-j-1}$. Other examples arise from any given positive constants a and b and completely monotone f on $[0, \infty)$ by setting

$$\mu_{ij} = f[j + (ja - j + 1)b].$$

THEOREM 1. *If (μ_{ij}) is a completely monotone matrix, then (μ_{ii}) is a completely monotone sequence.*

Proof:

$$\Delta^k \mu_{ii} = \sum_{j=0}^k \binom{k}{j} \Delta_1^j \Delta_2^{k-j} \mu_{i,i+j} \geq 0; \quad k = 0, 1, 2, \dots$$

Turning now toward the analogous theorem for two-place functions, we begin with a theorem [2, p. 9] which generalizes the classical Bernstein representation of completely monotone sequences by Riemann-Stieltjes integration with respect to a bounded nondecreasing integrator. The generalized theorem states that a necessary and sufficient condition for (1) to hold is that there exist a two-place distribution function Φ satisfying

$$(2) \quad \mu_{ij} = \int_0^1 \int_0^1 u^i v^j d\Phi; \quad i, j = 0, 1, 2, \dots$$

By a **two-place completely monotone function** we mean a function f on $[0, \infty)^2$, all of whose partial derivatives of all orders exist and satisfy

$$(3) \quad (-1)^{n+m} D_1^n D_2^m (f) \geq 0; \quad n, m = 0, 1, 2, \dots$$

Analogous to the representation (2), we have (3) if and only if there exists a two-place distribution function Φ satisfying

$$(4) \quad f(x, y) = \int_0^1 \int_0^1 u^x v^y d\Phi; \quad (x, y) \in [0, \infty)^2.$$

THEOREM 2. *If $f(x, y)$ is a two-place completely monotone function on $[0, \infty)^2$, then the function $h(x) = f(x, x)$ is completely monotone on $[0, \infty)$.*

Proof: From (4), we have

$$D_1^j D_2^{k-j} f(x, y) = \int_0^1 \int_0^1 u^x (\log u)^j v^y (\log v)^{k-j} d\Phi; \quad j = 0, 1, \dots, k.$$

The integral is clearly nonnegative for even k and nonpositive for odd k . Thus, from the identity

$$h^{(k)}(t) = \sum_{j=0}^k \binom{k}{j} D_1^j D_2^{k-j} f(x, y) \Big|_{x=t, y=t},$$

we conclude that $(-1)^k h^{(k)} \geq 0$; $k = 0, 1, 2, \dots$.

3. Derivative matrices. Suppose f is completely monotone on $[0, \infty)$. The rows of the matrix $(\mu_{ij}) = ((-1)^i f^{(i)}(j))$ are then completely monotone sequences [3, p. 164], but the columns generally are not completely monotone sequences.

THEOREM 3. *If f is completely monotone on $[0, \infty)$, then the sequence $((-1)^n f^{(n)}(n))$ is completely monotone.*

Proof. Writing $f(x) = \int_0^1 t^x d\alpha(t)$ (by Bernstein's theorem) and $\mu_n = (-1)^n f^{(n)}(n)$, we have for $k = 0, 1, 2, \dots$,

$$\begin{aligned}\Delta^k \mu_n &= (-1)^n \sum_{j=0}^k \binom{k}{j} f^{(n+j)}(n+j) \\ &= (-1)^n \int_0^1 (t \log t)^n (1 + t \log t)^k d\alpha(t) \geq 0.\end{aligned}$$

Example. Starting with $\alpha(t) = t$, we have in Theorem 3 the function $f(x) = 1/(x+1)$. By Theorem 3, the sequence $(n+1)^{-(n+1)} n!$ is completely monotone. One may conjecture and easily verify that the corresponding function $x^{-x} \Gamma(x)$ is also completely monotone.

References

1. R. P. Boas, Signs of Derivatives and Analytic Behavior, this MONTHLY, 78 (1971) 1085–1093.
2. J. A. Shohat and J. D. Tamarkin, The Problem of Moments, American Mathematical Society, Providence, 1950.
3. D. V. Widder, The Laplace Transform, Princeton University Press, Princeton, 1946.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

THE PERMANENT OF A DOUBLY STOCHASTIC MATRIX

RUSSELL MERRIS, California State University, Hayward

I. Statement of the problem. If $A = (a_{ij})$ is an n -square matrix, the permanent of A is the scalar-valued function of A defined by

$$\text{per}(A) = \sum a_{1i_1} a_{2i_2} \cdots a_{ni_n},$$

where the summation extends over all permutations i_1, i_2, \dots, i_n of the integers $1, 2, \dots, n$. Loosely speaking, the permanent is the determinant without the alternating minus signs. A good introduction to the permanent can be found in [8] or [12].

A matrix with nonnegative entries, and whose row and column sums are all equal to one, is called **doubly stochastic**. B. L. van der Waerden has conjectured [13] that $\text{per } A \geq n!/n^n$ for all doubly stochastic n -square matrices A , with equality holding if and only if $A = J_n$, the matrix each of whose entries is $1/n$.

Results of Marcus and Newman [10] have led to the following conjecture [4] which is stronger than van der Waerden's:

$$(1) \quad n^2 \text{per}(A) \geq \sum_{i,j=1}^n \text{per}(A_{ij})$$

for all doubly stochastic A , where A_{ij} is the submatrix of A obtained by deleting row i and column j . (Inequality (1) has been demonstrated for A positive semi-definite symmetric in [6] and [9].)

Consider now the **permanental adjoint** of the matrix A . It is the n -square matrix whose i, j entry is $\text{per}(A_{ji})$. Then (1) asserts that, on the average, $\text{per}(A)$ dominates the elements of the permanental adjoint of A .

If one could prove that $n \text{per}(A)$ dominated the maximum row sum of the permanental adjoint of A , then of course (1) would follow. Unfortunately, this is not the case. Let

$$A = \frac{1}{12} \begin{bmatrix} 0 & 7 & 5 \\ 7 & 0 & 5 \\ 5 & 5 & 2 \end{bmatrix}.$$

The permanental adjoint of A is

$$P = \frac{1}{144} \begin{bmatrix} 25 & 39 & 35 \\ 39 & 25 & 35 \\ 35 & 35 & 49 \end{bmatrix}.$$

The third row sum of P is $119/144$. Thrice the permanent of A is $112/144$.

Well, what about the minimum row sum? A *prima facie* weaker statement than (1) is that $n \text{per}(A)$ dominates the minimum row sum of the permanental adjoint of A , i.e.,

$$(2) \quad n \text{per}(A) \geq \min_i \sum_{j=1}^n \text{per}(A_{ji}).$$

This is the problem which I wish to propose: Can (2) be demonstrated for all doubly stochastic n -square matrices A ?

II. Equivalent statement of the problem. Let $M = (m_{ij})$ be any n -square matrix. Denote by $r_i(M)$ the i th row sum of M , and by $r(M)$ the sum of the elements of M . If A is any doubly stochastic matrix, then $r_i(M) = r_i(MA)$, for

$$r_i(MA) = \sum_{j=1}^n \sum_{k=1}^n m_{ik} a_{kj} = \sum_{k=1}^n m_{ik} \sum_{j=1}^n a_{kj} = \sum_{k=1}^n m_{ik} = r_i(M).$$

If P is the permanent adjoint of A , what does PA look like? By the Laplace expansion theorem (for permanents [7, p. 20]), the i, i element of PA is $\text{per}(A)$, for all i . But, whereas the product of A with its determinantal adjoint is $\det(A)I$, the off diagonal terms of PA are not zero. However, by our little argument above, conjecture (1) is equivalent to the statement that, on the average, $\text{per}(A)$ dominates the off diagonal elements of PA . More precisely, $n^2 \text{per}(A) \geq r(PA)$. Similarly, an equivalent formulation of (2) is

$$(3) \quad n \text{per}(A) \geq \min_i r_i(PA).$$

III. A method of attack. It is known that if N is a square matrix with non-negative entries, then the maximum eigenvalue of N dominates the minimum row sum of N [5, pp. 63 and 68]. Therefore, (2) would be established if one could prove that $n \text{per}(A)$ dominated the maximum eigenvalue of P . Alternatively, (3) would be proved if one could show that $n \text{per}(A)$ ($= \text{trace } PA$) dominated the maximum eigenvalue of PA .

References

1. R. A. Brualdi, Permanent of the product of doubly stochastic matrices, *Proc. Cambridge Philos. Soc.*, 62 (1966) 643-648, Lemma 1.
2. R. A. Brualdi and M. Newman, Inequalities for permanents and permanent minors, *Proc. Cambridge Philos. Soc.*, 61 (1965) 741-746.
3. ———, Inequalities for the permanent minors of nonnegative matrices, *Canad. J. Math.*, 18 (1966) 608-615.
4. D. Z. Dokovic, On a conjecture by van der Waerden, *Mat. Vesnik*, 4 (19) (1967) 272-276.
5. F. R. Gantmacher, *The theory of matrices*, vol. 2, Chelsea, New York, 1960.
6. Marvin Marcus and Russell Merris, A relation between the permanent and determinantal adjoints, *J. Australian Math. Soc.*
7. Marvin Marcus and Henryk Minc, *A survey of matrix theory and matrix inequalities*, Allyn and Bacon, Boston, 1964.
8. ———, Permanents, this MONTHLY, 72 (1965) 577-591.
9. ———, Extensions of classical matrix inequalities, *Linear Algebra and Appl.*, 1 (1968) 421-444.
10. Marvin Marcus and Morris Newman, On the minimum of the permanent of a doubly stochastic matrix, *Duke Math. J.*, 26 (1959) 61-72.
11. Henryk Minc, On lower bounds for permanents of (0,1) matrices, *Proc. Amer. Math. Soc.*, 22 (1969) 117-123, Lemma 1.
12. Herbert John Ryser, *Combinatorial mathematics*, Carus Monograph, No. 14, MAA, 1963.
13. B. L. van der Waerden, Aufgabe 56, *Jber. Deutsch. Math.-Verein*, 35 (1926) 117.

CLASSROOM NOTES

EDITED BY ROBERT GILMER

Material for this Department should be sent to David Roselle, Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803.

A GENERALIZATION OF A THEOREM OF ARCHIMEDES

WALTER RUDIN, University of Wisconsin

THEOREM A. *If two parallel planes whose distance is d intersect a sphere S of radius r , then the area of the part of S that lies between the two planes is $2\pi rd$.*

This is the theorem to which the title of this note alludes. (See p. 145 of C. B. Boyer's *A History of Mathematics*, Wiley, 1968.)

If the euclidean space R^3 is equipped with its usual coordinate system, so that the unit sphere S^2 is the set of all $x = (x_1, x_2, x_3)$ with $x_1^2 + x_2^2 + x_3^2 = 1$, then Theorem A is seen to be equivalent to

THEOREM A'. *If $0 \leq \delta \leq 1$ and if $E(\delta)$ is the set of all $x \in S^2$ with $|x_3| \leq \delta$, then the area of $E(\delta)$ is $4\pi\delta$.*

This formulation of Archimedes' theorem leads to a problem which may interest a good Calculus class when multiple integrals are studied.

Fix integers n and k , $1 \leq k \leq n-1$. Write each $x = (x_1, \dots, x_n) \in R^n$ in the form

$$(1) \quad x = (x', x''),$$

where $x' = (x_1, \dots, x_k) \in R^k$, $x'' = (x_{k+1}, \dots, x_n) \in R^{n-k}$. Define

$$(2) \quad \|x'\| = (x_1^2 + \dots + x_k^2)^{\frac{1}{2}}, \quad \|x''\| = (x_{k+1}^2 + \dots + x_n^2)^{\frac{1}{2}}.$$

The unit sphere S^{n-1} in R^n is then the set of all $x \in R^n$ with $\|x'\|^2 + \|x''\|^2 = 1$. For $0 \leq \delta \leq 1$, let $E(\delta)$ be the set of all $x \in S^{n-1}$ for which $\|x'\| \leq \delta$.

QUESTION. *For which pairs (n, k) is the $(n-1)$ -dimensional volume of $E(\delta)$ proportional to δ^{n-k} ?*

Theorem A' asserts that $(3, 2)$ is such a pair. The answer to the question has a feature which may be surprising: It depends only on k .

THEOREM 1. *The $(n-1)$ -dimensional volume of $E(\delta)$ is proportional to δ^{n-k} if and only if $k = 2$.*

The computation which proves Theorem 1 is made easy by Theorem 2 which will now be formulated. It reduces the computation of certain n -dimensional volumes to the evaluation of integrals over plane regions.

Suppose $1 \leq k \leq n-1$, as above. Let Q be the closed first quadrant in R^2 ; explicitly, Q is the set of all (ξ, η) with $\xi \geq 0$, $\eta \geq 0$. Let ϕ be the mapping of R^n onto Q defined by

$$(3) \quad \phi(x) = \phi(x', x'') = (\|x'\|, \|x''\|).$$

To see what ϕ does, observe that if $q = (\xi, \eta) \in Q$ then $\phi^{-1}(q)$ is the cartesian product of a $(k-1)$ -dimensional sphere and an $(n-k-1)$ -dimensional one, unless $\xi = 0$ or $\eta = 0$; when $\xi = 0 < \eta$, $\phi^{-1}(q)$ is a sphere of dimension $n-k-1$; when $\xi > 0 = \eta$, then $\phi^{-1}(q)$ is a sphere of dimension $k-1$; when $\xi = 0 = \eta$, $\phi^{-1}(q)$ is a point.

Let V_n be the n -dimensional volume of the unit ball in R^n . For example, $V_1 = 2$, $V_2 = \pi$. Let σ_{n-1} be the $(n-1)$ -dimensional volume of S^{n-1} . Thus $\sigma_0 = 2$, $\sigma_1 = 2\pi$, $\sigma_2 = 4\pi, \dots$. (The other values of V_n and σ_{n-1} are in (14) and (15).)

In general, let $m_n(A)$ denote the n -dimensional volume (or measure) of the set A .

THEOREM 2. *If Ω is a region in Q and if $A = \phi^{-1}(\Omega)$ is the set of all $x = (x', x'') \in R^n$ with $(\|x'\|, \|x''\|) \in \Omega$, then*

$$(4) \quad m_n(A) = k(n-k)V_k V_{n-k} \iint_{\Omega} \xi^{k-1} \eta^{n-k-1} d\xi d\eta.$$

Proof of Theorem 2. Let Ω first be a rectangle, given by $0 \leq a < \xi < b$, $0 \leq \alpha < \eta < \beta$. Then $A = A' \times A''$, where A' is the set of all $x' \in R^k$ with $a < \|x'\| < b$, and A'' is the set of all $x'' \in R^{n-k}$ with $\alpha < \|x''\| < \beta$. Hence

$$(5) \quad m_n(A) = m_k(A') m_{n-k}(A'') = (b^k - a^k) V_k \cdot (\beta^{n-k} - \alpha^{n-k}) V_{n-k}.$$

On the other hand,

$$(6) \quad \iint_{\Omega} \xi^{k-1} \eta^{n-k-1} d\xi d\eta = \frac{b^k - a^k}{k} \cdot \frac{\beta^{n-k} - \alpha^{n-k}}{n-k}.$$

Comparison of (5) and (6) shows that (4) holds for these rectangles. Hence (4) holds for general Ω , by any of the standard approximation procedures. (In fact, the collection of all sets Ω for which (4) holds is easily seen to be a σ -algebra, so that (4) holds for every Borel set Ω ; it also holds for every Lebesgue-measurable Ω .)

Proof of Theorem 1. Let $C(\delta)$ be the cone with base $E(\delta)$ and vertex at the origin. In other words, $C(\delta)$ is the union of all intervals in R^n which have one endpoint at the origin and the other in E . Or, $C(\delta) = \{tx: 0 \leq t \leq 1, x \in E(\delta)\}$. For $r > 0$, note that

$$(7) \quad m_n(rC(\delta)) = r^n m_n(C(\delta)),$$

and that $m_{n-1}(E(\delta))$ is the derivative of the left side of (7), evaluated at $r = 1$. Since the derivative of r^n is n when $r = 1$, (7) implies that

$$(8) \quad m_{n-1}(E(\delta)) = nm_n(C(\delta)).$$

The special case $\delta = 1$ yields

$$(9) \quad \sigma_{n-1} = nV_n.$$

Note also that $C(\delta) = \phi^{-1}(\Omega)$, where Ω consists of all $(\xi, \eta) \in Q$ that satisfy

$$(10) \quad \xi^2 + \eta^2 \leq 1 \text{ and } \eta \leq (\tan \alpha) \cdot \xi;$$

here α is chosen so that $\sin \alpha = \delta$, $0 \leq \alpha \leq \pi/2$.

By switching to polar coordinates and then setting $t = \sin \theta$, it now follows from (8), (9), and Theorem 2 that

$$\begin{aligned} m_{n-1}(E(\delta)) &= n\sigma_{k-1}\sigma_{n-k-1} \iint_{\Omega} \xi^{k-1}\eta^{n-k-1} d\xi d\eta \\ (11) \quad &= n\sigma_{k-1}\sigma_{n-k-1} \int_0^1 r^{k-1} r^{n-k-1} r dr \int_0^\alpha (\cos \theta)^{k-1} (\sin \theta)^{n-k-1} d\theta \\ &= \sigma_{k-1}\sigma_{n-k-1} \int_0^\delta (1-t^2)^{(k-2)/2} t^{n-k-1} dt. \end{aligned}$$

It is now clear that $m_{n-1}(E(\delta))$ is proportional to δ^{n-k} if and only if the last integrand is proportional to t^{n-k-1} , and this happens if and only if $k = 2$.

Theorem 1 is thus proved.

REMARKS. When $k = 2$, (11) and (9) yield

$$(12) \quad m_{n-1}(E(\delta)) = \sigma_1\sigma_{n-3} \cdot \frac{\delta^{n-2}}{n-2} = 2\pi V_{n-2} \delta^{n-2}$$

which reduces to Theorem A' when $n = 3$. If $\delta = 1$, (12) and (9) give the recursion formula

$$(13) \quad nV_n = 2\pi V_{n-2}.$$

Since $V_1 = 2$ and $V_2 = \pi$, (13) enables us to compute V_n for all n . By induction,

$$(14) \quad V_n = \frac{2}{n} \cdot \frac{\pi^{n/2}}{\Gamma(n/2)}.$$

(The only properties of the gamma function that are needed here are: $\Gamma(1) = 1$, $x\Gamma(x) = \Gamma(x+1)$, $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.)

Finally, (9) and (14) give

$$(15) \quad \sigma_{n-1} = \frac{2\pi^{n/2}}{\Gamma(n/2)}.$$

This research was partially supported by NSF Grant GP-24182.

THE CHROMATIC POLYNOMIAL OF A COMPLETE BIPARTITE GRAPH

J. R. SWENSON, University of Toronto

We use [1] for all definitions in graph theory. In particular, we only consider finite undirected graphs without loops or multiple edges. A **bipartite graph** is one whose point set can be partitioned into two sets, U and V , such that every edge of the graph runs from U to V . It is complete if every point of U is connected to every point of V [1, p. 17]. If there are p points in U and q points in V , then we denote the complete bipartite graph by $K_{p,q}$. Denote by $f(G, t)$ the number of different colorings of the labelled graph G using t colors; $f(G, t)$ is a polynomial called the **chromatic polynomial** of G [1, pp. 146 ff].

The chromatic polynomial of a graph has some general properties which are easy to determine, e.g., its coefficients alternate in sign [1, p. 147], and there is a general formula for $f(G, t)$. The latter however, is difficult to evaluate in general. We offer here an argument for evaluating $f(K_{p,q}, t)$.

Let $K_{p,q}$ be partitioned into its sets U and V and let there be t colors. Choose r colors. Let there be $c(p, t, r)$ ways to color the p independent vertices of U using r colors chosen from the t colors. Then exactly $t - r$ colors remain and these may be used in exactly $(t - r)^q$ ways to color the remaining vertices in V . Thus there are

$$\sum_{r=1}^p c(p, t, r)(t - r)^q$$

ways to color $K_{p,q}$. Our problem is to evaluate $c(p, t, r)$.

There are $\binom{t}{r}$ ways to choose r colors from t colors. The number of ways to color the p distinguishable points with exactly these r colors is equivalent to distributing p distinguishable balls into r distinguishable boxes leaving no box empty. By Riordan [2, p. 91] the latter can be done in exactly

$$r!S(p, r)$$

ways, where $S(p, r)$ is a Stirling number of the second kind. Thus,

$$c(p, t, r) = r!S(p, r) \binom{t}{r}.$$

We can simplify

$$(1) \quad \sum_{r=1}^p \binom{t}{r} r!S(p, r)(t - r)^q$$

as follows. Let $(t)_r = (t)(t - 1) \cdots (t - r + 1)$ be the "falling factorial". Then

$$\binom{t}{r} r! = (t)_r.$$

Also, from Riordan [2, p. 33] we have $(t-r)^q = \sum_{s=1}^q S(q,s)(t-r)_s$. As $(t)_r(t-r)_s = (t)_{r+s}$ we can obtain after substitution in (1)

$$f(K_{p,q}, t) = \sum_{r=1}^p \sum_{s=1}^q S(p,r)S(q,s)(t)_{r+s}.$$

The last expression shows the symmetry in p and q which must obtain since $K_{p,q} \cong K_{q,p}$. It also shows that the polynomial is of degree $p+q$ and, as each expression $(t)_{r+s}$ is monic and $S(p,p) = S(q,q) = 1$, we have that the leading coefficient of f is 1.

References

1. F. Harary, Graph Theory, Addison-Wesley, Reading, Mass., 1969.
2. J. Riordan, An Introduction to Combinatorial Analysis, Wiley, New York, 1958.

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to Shirley Hill, Department of Mathematics, University of Missouri, Kansas City, MO 64110, or to Paul Mielke, Department of Mathematics, Wabash College, Crawfordsville, IN 47933.

TRAINING SECONDARY MATHEMATICS TEACHERS IN VENEZUELA

D. B. AICHELE, Oklahoma State University

As a visiting professor of mathematics at the Universidad de Carabobo in Valencia, Venezuela, I had the opportunity to travel quite extensively throughout the country. Since my primary duties at Oklahoma State University are in the area of teacher training, I naturally took advantage of opportunities to learn of the Venezuelan approach to training secondary mathematics teachers. In this regard, this paper summarizes the policies of the government-supported Instituto Pedagógico Experimental de Barquisimeto, which is one of the eight institutions (6 government- and 2 privately-supported) charged with preparing secondary level teachers. Its program is fairly representative of secondary level teacher preparation conducted in Venezuela.

Before looking at the actual teacher education program at Barquisimeto, I believe it is necessary to understand something of Venezuelan education in general. Although Venezuela is perhaps the richest and most educated of all the Latin American countries, it nonetheless has a literacy problem. In the fight against illiteracy, 1.5 million adults have been taught how to read and write during the past 10 years; but still 11% of the 10.4 million Venezuelans are illiterate.

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

ASSOCIATE EDITORS: JOSHUA BARLAZ, ERIC S. LANGFORD. COLLABORATING EDITORS: LEONARD CARLITZ, GULBANK D. CHAKERIAN, HASKELL COHEN, S. ASHBY FOOTE, ISRAEL N. HERSTEIN, MURRAY S. KLAMKIN, DANIEL J. KLEITMAN, ROGER C. LYNDON, MARVIN MARCUS, CHRISTOPH NEUGEBAUER, ALBERT WILANSKY, AND UNIVERSITY OF MAINE PROBLEMS GROUP: EARL M. L. BEARD, GEORGE S. CUNNINGHAM, CLAYTON W. DODGE, OSKAR FEICHTINGER, WILLIAM R. GEIGER, RAMESH GUPTA, GARY HAGGARD, PHILIP M. LOCKE, JOHN C. MAIRHUBER, CURTIS S. MORSE, GRATAN P. MURPHY, EDWARD S. NORTHAM AND WILLIAM L. SOULE, JR.

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of elementary Problems in this issue should be typed (with double spacing) and should be mailed before December 31, 1973.

An asterisk () means neither the proposer nor the editors supplied a solution.*

E 2426*. *Proposed by C. A. Long, Bowling Green University*

It is easy to show that the equilateral triangle can be inscribed in a square, and that a square can be inscribed in a regular pentagon. Can a regular pentagon be inscribed in a regular hexagon?

E 2427*. *Proposed by Harry Ruderman, Hunter College Campus School*

Suppose that 1 is written as a sum of n distinct Egyptian fractions. Find upper and lower bounds for the smallest fraction in the sum.

E 2428. *Proposed by M. S. Klamkin, Ford Motor Company*

If a_i ($i = 1, 2, \dots, n$) denote real numbers, show that

$$n \min(a_i) \leq \sum a_i - S \leq \sum a_i + S \leq n \max(a_i),$$

where

$$(n-1)S^2 = \sum_{1 \leq i < j \leq n} (a_i - a_j)^2 \quad (S \geq 0)$$

and with equality if and only if $a_i = \text{constant}$.

E 2429. *Proposed by E. T. H. Wang, University of British Columbia*

Let k_n denote the least integer such that every $n \times n$ matrix of zeros and ones with exactly k_n ones in each row and in each column contains a 2×2 submatrix without zero. Obtain a lower estimate for k_n and discuss the case of equality.

E 2430. *Proposed by John Masley, University of Illinois at Chicago Circle*

Let a and m denote natural numbers, and let ϕ denote Euler's totient function. Euler's generalization of Fermat's "Little Theorem" asserts that if $(a, m) = 1$, then

$$(*) \quad a^{\phi(m)+1} \equiv a \pmod{m}.$$

Show that $(*)$ holds if and only if the following: if p is a prime that divides a , then $p^k \mid a$ whenever $p^k \mid m$.

E 2431. *Proposed by Alan McConnell, Howard University*

Consider a finite field F with elements a_1, a_2, \dots, a_n . Form the Vandermonde matrix $V(a_1, \dots, a_n) = (v_{ij})$, where $v_{ij} = (a_j)^{i-1}$ for $i, j = 1, 2, \dots, n$ (and where $0^0 = 1$). Find V^{-1} and evaluate $\det V$ (where the operations are carried out in F).

SOLUTIONS OF ELEMENTARY PROBLEMS

A Quadrilateral Proportion

E 1085 [1953, 551]. *Proposed by Josef Langr, Prague, Czechoslovakia*

The perpendicular bisectors of the sides of a quadrilateral Q form a quadrilateral Q_1 , and the perpendicular bisectors of the sides of Q_1 form a quadrilateral Q_2 . Show that Q_2 is similar to Q and find the ratio of similitude.

Partial Solution by Martin Thomas, Shirley, Southampton, England. Let $ABCD$ be Q . Let A', B', C', D' be the circumcenters of BCD, CDA, DAB, ABC , so that $A'B'C'D'$ is Q_1 . Let $A''B''C''D''$ be Q_2 . Now A' and C' lie on the perpendicular bisector of BD , and similarly B'' and D'' lie on the perpendicular bisector of $A'C'$. Thus $B''D''$ and BD are parallel. Also AB and $A''B''$, BC and $B''C''$, etc., are pairs of parallel lines. Hence triangles ABD and $A''B''D''$ are directly similar, as are also BCD and $B''C''D''$. It follows that $ABCD$ and $A''B''C''D''$ are similar.

To see that the ratio of similitude can be any positive real number, note that if Q is a rhombus whose acute angle is θ , then Q_1 is a similar rhombus rotated through 90° . If their ratio of similitude is r , then $r \rightarrow 0$ as $\theta \rightarrow 90^\circ$ and $r \rightarrow \infty$ as $\theta \rightarrow 0^\circ$. For $\theta = 45^\circ$, $r = 1$ and Q and Q_2 coincide.

In the general case an expression for the ratio of similitude of Q_2 to Q seems prohibitively involved. [A "reasonable" expression for this ratio is solicited — Ed.]

A Difficult Triangle Inequality

E 2245 [1970, 652; 1971, 793; 1972, 1034]. *Proposed by A. W. Walker, Toronto, Canada*

If A, B, C ; a, b, c ; s are the angles, side lengths, and semi-perimeter of any plane triangle, then

$$(a + b + c)^3(s - a)(s - b)(s - c) \geq (a^2 + b^2 + c^2)^3 \cos A \cos B \cos C.$$

III. *Comment by Robert Breusch, Amherst College.* We shall show how the following inequality of Van Tooren [1972, 1035, lines 7-8] (equivalent to Walker's inequality above) does indeed hold for all nonnegative a, b, c :

$$(a + b + c)(-a + b + c)(a - b + c)(a + b - c)[(abc)^2(a + b + c)^2 - (a^2 + b^2 + c^2)^4] + 8(abc)^2(a^2 + b^2 + c^2)^3 \geq 0.$$

Assume that a, b, c are nonnegative and (without loss of generality) that $a + b + c = 1$. Let $x = a^2 + b^2 + c^2$ and $y = abc$. With a little rearrangement Van Tooren's inequality becomes

$$K(x, y) = 8x^3y^2 + (x^4 - y^2)(2x + 8y - 1) \geq 0.$$

Clearly $1/3 \leq x \leq 1$ and $0 \leq y \leq 1/27$, so that $x^4 - y^2 > 0$ and $K(x, y) > 0$ whenever $1/2 < x \leq 1$; we can thus restrict our attention to those x which satisfy $1/3 \leq x \leq 1/2$. Now $ab + bc + ca = (1 - x)/2$ so that a, b, c are the three zeros of $f(t) = t^3 - t^2 + \frac{1}{2}(1 - x)t - y$ and it is known that the polynomial $t^3 - At^2 + Bt - C$ has three real zeros if and only if

$$4B^3 - A^2B^2 + 4A^3C - 18ABC + 27C^2 \leq 0.$$

It follows that x and y must satisfy an inequality which, after some rearrangement, can be written in the form

$$\left(y - \frac{5 - 9x}{54}\right)^2 \leq \frac{(6x - 2)^3}{108^2},$$

that is, $y \geq m(x)$, where

$$m(x) = \frac{5 - 9x}{54} - \frac{(6x - 2)^{3/2}}{108}.$$

Calculating the partial derivative, we see that

$$\frac{\partial K}{\partial y} = 8(x^4 - 3y^2) + 2y(8x^3 - 2x + 1),$$

which is positive since $x \geq 1/3$ and $y \leq 1/27$. This means that $K(x, y) \geq K(x, m(x))$

$\equiv F(x)$, so that if we can show that $F(x) \geq 0$ for $1/3 \leq x \leq 1/2$, we are done. Making the transformation $z = t(x) = (6x-2)^{1/2}$ so that $x = t^{-1}(z) = (z^2 + 2)/6$, we see that $0 \leq z \leq 1$ as $1/3 \leq x \leq 1/2$, and that

$$G(z) \equiv F(t^{-1}(z)) = 2^{-4}3^{-9}z^2(1-z)^5(72 + 32z + 49z^2 + 3z^3 + 7z^4 - z^5),$$

which is clearly nonnegative for $0 \leq z \leq 1$.

Iterated Composition of a Function of Prime Factors of n

E 2356 [1972, 518]. *Proposed by J. B. Roberts, Reed College*

If n is a natural number, define $f(n)$ to be 1 plus the sum of the prime factors of n , each prime being counted according to its multiplicity. For example, $f(12) = 8$. Prove that if n is greater than 6, then the sequence of iterates $n, f(n), f(f(n)), \dots$ contains an 8, and hence from some point on, must repeat: 8, 7, 8, 7, \dots .

Solution by Hans Kappus, Switzerland. Since $f(n) \geq 7$ for $n > 6$ and $f(7) = 8$, it suffices to prove the following property of $f(n)$:

If $n \geq 9$, then either n is composite and $f(n) \leq n - 2$, or n is prime whence $f(n) = n + 1$ (which is composite) and then $f(f(n)) \leq f(n) - 2 = n - 1$.

In fact, this is true for $n = 9$, so let us assume that it is true for all k such that $9 \leq k < n$. Let n be composite, $n = k_1 \cdot k_2$. Then $(k_1 - 1)(k_2 - 1) \geq 4$. Furthermore, $f(k_i) \leq k_i + 1$ if either $k_i \leq 6$ or k_i is prime ≥ 7 ; otherwise $f(k_i) \leq k_i - 2$ by assumption. Also

$$f(n) = f(k_1) + f(k_2) - 1.$$

Hence $f(n) \leq k_1 + 1 + k_2 + 1 - 1$. But

$$k_1 + 1 + k_2 + 1 - 1 = n + 2 - (k_1 - 1)(k_2 - 1).$$

Therefore $f(n) \leq n + 2 - 4 \leq n - 2$.

Also solved by Anders Bager (Denmark), S. Baskaran (India), Problem Solving Group of Berne (Switzerland), D. M. Bloom, Peter Bundschuh (Germany), R. J. Evans, Scott Forrest, Ray Glenn, Michael Goldberg, M. G. Greening (Australia), C. V. Heuer, W. M. Hill and his Linear Algebra Class, Wells Johnson, Václav Konečný, L. Kuipers, O. P. Lossers (Netherlands), Kevin McAvaney, Carolyn MacDonald, William Margolis, Helen M. Marston, Norman Miller, L. R. Nyhoff, M. R. Railkar (India), Eric Rosenthal, Steven Russ, Harry Sherman, Nan-Shan Shou, Allen Stenger, Walter Stromquist, R. K. Tamaki, S. J. Tillman, Mike Vitale, Charles Wexler, and the proposer.

The Non-disjointness of Infinitely Many Sets having the Same Probability

E 2362 [1972, 663]. *Proposed by C. H. Kimberling, University of Evansville*

Suppose that in some probability space, E_1, E_2, \dots are events with common probability p . Let $m \geq 2$ be a fixed integer. Prove or disprove that

$$p^m \leq \sup \{P(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_m})\},$$

where the supremum is taken over all m -tuples (i_1, i_2, \dots, i_m) of distinct natural numbers.

Solution by Ellen Hertz, Bronx, N.Y. Let X_i be the indicator variable of the event E_i , $i = 1, 2, \dots$. By Jensen's inequality [see Parzen, *Modern Probability Theory and its Applications*, p. 434]

$$E(((X_1 + \cdots + X_n)/n)^m) \geq p^m.$$

If we expand $(X_1 + \cdots + X_n)^m$ without collecting terms we obtain a sum of n monomials of which $n(n-1)\cdots(n-m+1)$ contain m distinct X_i . Write $(X_1 + \cdots + X_n)^m = T_1 + T_2$ where T_2 consists of the products of m distinct X_i and T_1 consists of those with fewer. $E(T_1)$ is the sum of $n^m - n(n-1)\cdots(n-m+1)$ nonnegative terms each at most p so that

$$0 \leq E(T_1)/n^m \leq (n^m - n(n-1)\cdots(n-m+1))p/n^m = o(1) \text{ as } n \rightarrow \infty.$$

Let S_n be the set of all m -tuples (i_1, \dots, i_m) of distinct natural numbers such that $1 \leq i_j \leq n$, $j = 1, 2, \dots, m$. Then

$$\begin{aligned} E(T_2) &= \sum_{(i_1, \dots, i_m) \in S_n} E(X_{i_1} \cdots X_{i_m}) \\ &\leq n(n-1)\cdots(n-m+1) \sup_{(i_1, \dots, i_m) \in S_n} E(X_{i_1} \cdots X_{i_m}). \end{aligned}$$

It follows that

$$p^m \leq (E(T_1) + E(T_2))/n^m \leq o(1) + \sup_{(i_1, \dots, i_m) \in S_n} E(X_{i_1} \cdots X_{i_m}).$$

The conclusion now follows upon taking the limit as $n \rightarrow \infty$.

Also solved by Janos Galambos, J. Gillis (Israel), J. C. Kieffer, Gérard Letac (France), and Andrew Odlyzko.

Editorial Note. Gillis and Odlyzko give extensions based on an infinite version of Ramsey's theorem. Galambos cites J. Galambos and A. Renyi, *On quadratic inequalities in the theory of probability*, *Studia Scient. Math. Hungar.*, 3 (1968), 351-358 and J. Galambos, *Quadratic inequalities among probabilities*, *Ann. Univ. Sci. Budapest, Eötvös, Sect. Math.*, 12 (1969), 11-16 which contain related results.

An Integer Inequality

E 2368 [1972, 772]. Proposed by C. V. Heuer, Concordia College

Prove that if $1 < x_1 < x_2 < \cdots < x_k < y_1 < y_2 < \cdots < y_m$ are integers such that $\sum x_i \geq \sum y_i$, then $\prod x_i > \prod y_i$.

Solution by Ivan Niven, University of Oregon. It is convenient to prove a little more, namely that the same conclusion holds if

$$1 < x_1 < x_2 \leq x_3 \leq x_4 \leq \cdots \leq x_k < y_1 \leq y_2 \leq \cdots \leq y_m.$$

Note that $k \geq 2$ and $k > m$. Letting $s = \sum x_i$, the proof is by induction on k and s . For $k = 2$ we note that $m = 1$, $x_1 \geq 2$, $x_2 \geq 3$, $s \geq 5$. Then $x_1 + x_2 \geq y_1$ implies $x_1 x_2 > y_1$ because $x_1 x_2 - x_1 - x_2 = (x_1 - 1)(x_2 - 1) - 1 > 0$. This proves the result for $k = 2$, so now suppose the result holds for all $k < j$ and look at the case $k = j$. To get a start on induction on s , we note that the least value of s is $3j - 1$ from the values $x_1 = 2$, $x_2 = x_3 = \cdots = x_j = 3$. Here $\prod x_i = 2 \cdot 3^{j-1}$ and $y_i \geq 4$ for all i . By the inequality of the arithmetic-geometric mean,

$$\prod y_i \leq (\sum y_i / m)^m \leq (s/m)^m = ((3j - 1)/m)^m.$$

Also $s/m \geq \sum y_i / m \geq 4$ and $m \leq s/4 = (3j - 1)/4$. Now with m so restricted, we prove easily that $\{(3j - 1)/m\}^m$ is an increasing function of m by differentiating its logarithm with respect to m to get $\log(3j - 1) - 1 - \log m$. This derivative is positive because $s/m \geq 4 > e$ gives $3j - 1 > me$. Hence it suffices to prove $2 \cdot 3^{j-1} > \{(3j - 1)/m\}^m$ for the largest possible m , namely $m = (3j - 1)/4$. Simple algebra completes this proof, and so the result holds for $k = j$ and $s = 3j - 1$.

Using induction on s with $k = j$, we look at some particular $s > 3j - 1$, assuming the result for smaller values of s . Note that $x_1 \geq 3$ or $x_j \geq 4$ in all cases.

In case $x_1 \geq 3$ we note that $x_1 - 1, x_2, x_3, \dots, x_j, y_1, y_2, \dots, y_{m-1}, y_m - 1$ form a set of values satisfying the hypotheses (with $y_m - 1$ in perhaps an earlier position in the ordering of the y 's), and likewise the set $x_2, x_3, \dots, x_j, y_1, y_2, \dots, y_{m-1}$. Applying the result in these two cases and adding, the proof is complete.

In case $x_1 = 2$, then $x_j \geq 4$ and $y_i \geq 5$ for $i = 1, 2, \dots, m$. We note that the sets $x_1, x_2, \dots, x_{j-1}, x_j - 1, y_1, y_2, \dots, y_{m-1}, y_m - 1$ and $x_1, x_2, \dots, x_{j-1}, y_1, y_2, \dots, y_m - 1$ satisfy the hypotheses, and these give inequalities on products which again add to give the desired result.

REMARK. The result is false if the x 's and y 's are assumed to be real numbers not necessarily integers, for example in the case $x_1 = 2, x_2 = 2.01, x_3 = 2.1, y_1 = 3, y_2 = 3.1$. However, under the assumptions $e \leq x_1 \leq x_2 \leq \cdots \leq x_k < y_1 \leq y_2 \leq y_3 \leq \cdots \leq y_m$ and $\sum x_i \geq \sum y_i$, then $\prod x_i > \prod y_i$ for real numbers x_i and y_i .

Also solved by Robert Breusch, Jordi Dou (Spain), Harry Lass, O. P. Lossers (Netherlands), Carolyn MacDonald, L. E. Mattics, Leo Ringwald, St. Olaf College Students, Wolfe Snow, Oto Strauch (Czechoslovakia), Jim Tattersall, Temple University Problem Solving Group, Louis Thurston, Phil Tracy & Joe Mercer, P. H. Young, Alexandras Zujus, and the proposer.

Just a Short (Random) Walk

E 2369 [1972, 773]. Proposed by Harry Lass, California Institute of Technology

For the two-dimensional symmetric random walk starting at the origin, show that the probability of reaching the point $(1, 0)$ before reaching any other point on the line $x = 1$, is $1 - 2/\pi$.

Solution by Frederick Carty, Akron, Ohio. Let S_m be the set of all paths of m steps in length starting at the origin, ending at $(1, 0)$ and not reaching the line $x = 1$ in the first $m - 1$ steps. Let A_m be the number of paths in S_m . Note that no path starting at the origin can reach $(1, 0)$ in an even number of steps, i.e., S_{2k} is empty and $A_{2k} = 0$.

To evaluate A_{2n-1} , we first find the number of paths in S_{2n-1} with $2j$ vertical steps. The number of ways the vertical steps can be arranged is $\binom{2j}{j}$. The number of ways the first $2n - 2j - 2$ horizontal steps can be arranged with no step to the right of the origin is $\binom{2n-2j-2}{n-j-1} (n-j)^{-1}$ as shown in Feller, *Introduction to Probability Theory and Its Applications*. The number of ways an arrangement of $2j$ vertical steps can be interspersed with an arrangement of $2n - 2j - 2$ horizontal steps is $\binom{2n-2}{2j}$. Thus

$$\begin{aligned} A_{2n-1} &= \sum_{j=0}^{n-1} \binom{2n-2}{2j} \binom{2j}{j} \binom{2n-2j-2}{n-j-1} \frac{1}{n-j} \\ &= \sum_{j=0}^{n-1} \binom{2n-2}{n-1} \frac{1}{n} \binom{n}{j} \binom{n-1}{j} \\ &= \binom{2n-1}{n} \frac{1}{2n-1} \sum_{j=0}^{n-1} \binom{n}{j} \binom{n-1}{j} \\ &= \binom{2n-1}{n}^2 \frac{1}{2n-1}. \end{aligned}$$

Finally the desired probability p_0 is given by:

$$\begin{aligned} p_0 &= \sum_{n=1}^{\infty} A_{2n-1} 4^{-2n+1} = \sum_{n=1}^{\infty} \binom{2n-1}{n}^2 \frac{1}{2n-1} 4^{-2n+1} \\ &= \sum_{n=1}^{\infty} \left(\frac{1 \cdot 3 \cdot \dots \cdot (2n-1)}{2 \cdot 4 \cdot \dots \cdot 2n} \right)^2 \frac{1}{2n-1}. \end{aligned}$$

This final sum equals $1 - 2/\pi$ as found in Jolley, *Summation of Series*, p. 73.

Also solved by Ellen Hertz and by the proposer.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers—The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before December 31, 1973. Contributors (in the United States) who desire acknowledgment of receipt of their solutions are asked to enclose self-addressed, stamped postcards.

An asterisk () means neither the proposer nor the editors supplied a solution.*

5922. *Proposed by Paul Cohn, Bedford College, London, England*

A and B are two $m \times n$ matrices over an infinite field k such that $\text{rank}(A + \lambda B) \leq \text{rank } A$ for all $\lambda \in k$. Find P and Q of orders m, n respectively such that

$$B = PA + AQ, \quad PAQ = 0.$$

5923*. *Proposed by Emeric Deutsch, Polytechnic Institute of Brooklyn*

Let $\|\cdot\|$ be a norm on the vector space C^n of all n -tuples of complex numbers, and let A be an operator on C^n such that $\|(A - \alpha I)^{-1}\| = r[(A - \alpha I)^{-1}]$ for each complex α which is not in the spectrum of A (r denotes spectral radius). Is $\|A\| = r(A)$?

5924. *Proposed by Donald Girod, Canisius College, Buffalo, N.Y.*

A standard exercise in an introductory algebra course is to show that no group G can ever be the set theoretic union of two proper subgroups H_1, H_2 . It is possible, however, for a group to be the union of some finite number (> 2) of proper subgroups. For example, $\mathbb{Z} \oplus \mathbb{Z}$ is the union of three proper subgroups. Characterize those Abelian groups G having a finite set of proper subgroups $\{H_1, \dots, H_n\}$ such that $G = H_1 \cup \dots \cup H_n$.

5925. *Proposed by A. G. O'Farrell, Brown University*

Show that the matrix (a_{ij}) , where $a_{ij} = 1/(1 + |j - i|)$, is positive definite.

5926. *Proposed by R. P. Boas, Northwestern University*

If f and g are nonnegative, bounded, and integrable over $(-\infty, \infty)$, does it follow that

$$\int_{-\infty}^{\infty} \sup_t [g(x-t)f(t)]dx \geq \sup_x \int_{-\infty}^{\infty} g(x-t)f(t)dt?$$

5927. *Proposed by C. R. Johnson, National Bureau of Standards*

Find all convex subsets K of the complex plane C such that if L is any convex subset of C , then $\{zw: z \in K, w \in L\}$ is convex.

SOLUTIONS OF ADVANCED PROBLEMS

Compactness and Open Covers

5850 [1972, 399]. *Proposed by R. K. Tamaki, California State College at Los Angeles*

Let X be metrizable. Prove that X is compact if and only if, for every metric d for X , every open cover $\{U_\alpha\}$ of X has a Lebesgue number $\lambda > 0$ (i.e., we require that each d -ball $B_d(x, \lambda)$ is contained in some U_α).

Solution by A. A. Jagers, Twente University of Technology, Netherlands. Suppose first that X is compact. Fix a metric d for X and, given an open cover $\{U_\alpha\}$ of X , choose for each $x \in X$ a number $\delta(x) > 0$ such that $B_d(x, 2\delta(x)) \subset U_\alpha$ for some α . This yields an open cover $\{B_d(x, \delta(x))\}$ which contains, in view of the compactness of X , a finite subcover $\{B_d(x_i, \delta(x_i)): 1 \leq i \leq n\}$. Now one easily checks that $\lambda = \min\{\delta(x_i): 1 \leq i \leq n\}$ is a positive Lebesgue number for $\{U_\alpha\}$.

Conversely suppose that, for every metric d for X , every open cover $\{U_\alpha\}$ has a Lebesgue number $\lambda > 0$. Then, X being metrizable, it suffices to show that X is sequentially compact. To do so, suppose that $(a_n)_{n=1}^\infty$ is a sequence of points in X without convergent subsequence and with $a_n \neq a_m$ for $n \neq m$. Then any subset of $A = \{a_n: n = 1, 2, \dots\}$ is closed. Now two possibilities arise. (i) But for a finite number, all a_n are clusterpoints of X . (ii) There exists an infinite subsequence $(b_n)_{n=1}^\infty$ of $(a_n)_{n=1}^\infty$ such that all b_n are isolated points of X . In the first case there exists for each n a real number $\varepsilon_n > 0$ such that $\lim_{n \rightarrow \infty} \varepsilon_n = 0$, $B_d(a_n, 2\varepsilon_n) \not\subset B_d(a_n, \varepsilon_n)$ and $a_m \notin B_d(a_n, \varepsilon_n)$ for $m \neq n$ and it follows at once that $\{B_d(a_n, \varepsilon_n): n = 1, 2, \dots\} \cup \{X \setminus A\}$ is an open cover with no positive Lebesgue number. A contradiction. In the second case, set $B = \{b_n: n = 1, 2, \dots\}$ and consider a new metric d_1 given by $d_1(x, y) = d(x, y) [1 + d(x, y)]^{-1}$ if $x, y \notin B$, $d_1(x, b_n) = d_1(b_n, x) = 1$ if $x \notin B$, and $d_1(b_n, b_m) = 2|n^{-1} - m^{-1}|$. Then d_1 is also a metric for X and this time $\{B_d(b_n, n^{-2}): n = 1, 2, \dots\} \cup \{X \setminus B\}$ is an open cover of X with no positive Lebesgue number. This second contradiction completes the proof.

Also solved by O. P. Lossers (Netherlands), Simeon Reich (Israel), and the proposer. A number of incomplete arguments were received.

Translates of Sequences and the Cantor Set

5851 [1972, 399]. *Proposed by Douglas Lind, Stanford University*

Is there a bounded sequence of real numbers each translate of which has only finitely many terms in the Cantor set?

Solution by Don Coppersmith, Massachusetts Institute of Technology. Such sequences do exist. We prove that no translate of the sequence $x_n = 13/3^n$ has more than two elements of the Cantor set.

Recall that the Cantor set K is the set of numbers between 0 and 1, inclusive, which have ternary representations consisting entirely of 0's and 2's. Notice that $1/3 = (0.10000\ldots)_3 = (0.02222\ldots)_3$ is an element of the Cantor set. For consistency we shall always choose the first representation (that with a string of 0's) when presented with such an ambiguity. This is the only instance where a 1 can occur in the ternary representation; it must be followed by all 0's.

Given y , assume that there are three values of n , $L < M < N$, such that $y + x_n \in K$. Consider the difference between $(y + x_L)$ and $(y + x_N)$, two members of K . Depending on the value of $N - L$, this difference is one of the following (all arithmetic in the base 3):

TABLE I

111	111	1110...0
-111	-111	-0...0111
<hr/> 0222	<hr/> 10212	<hr/> 1102...2112

Here "0...0" represents a string of 0's of arbitrary positive or zero length, similarly for "2...2". The differences are all to be divided by 3^N , i.e., the "decimal" point is to be inserted.

These numbers can be differences between elements of K in only the following ways:

TABLE II

$xx0000yy$ $+0222$ <hr/> $xx0222yy$	$xx0001zz$ $+0222$ <hr/> $xx1000zz$	$xx2000yy$ $+0222$ <hr/> $xx2222yy$	$xx02\ldots2001zz$ $+0222$ <hr/> $xx10\ldots0000zz$
$xx02020yy$ $+10212$ <hr/> $xx20002yy$	$xx02021zz$ $+10212$ <hr/> $xx20010zz$	$xx02220yy$ $+10212$ <hr/> $xx20202yy$	$xx02221zz$ $+10212$ <hr/> $xx20210zz$
$xx020\ 0\ldots0020yy$ $+110\ 2\ldots2112$ <hr/> $xx200\ 2\ldots2202yy$	$xx022\ 0\ldots0020yy$ $+110\ 2\ldots2112$ <hr/> $xx202\ 2\ldots2202yy$	$xx020\ 0\ldots0021zz$ $+110\ 2\ldots2112$ <hr/> $xx200\ 2\ldots2210zz$	
	$xx022\ 0\ldots0021zz$ $+110\ 2\ldots2112$ <hr/> $xx202\ 2\ldots2210zz$		

Here xx and yy represent arbitrary strings of 0's and 2's; zz represents an infinite string of 0's; $0\ldots0$ represents a finite (possibly null) string of 0's, and similarly for $2\ldots2$. The top number in each case represents $y + x_N$ and the bottom is $y + x_L$. The rightmost explicit digit in the top number before the yy or zz is in the N th ternary place.

Now do the same for the pair $(y + x_N)$ and $(y + x_M)$. We get another addition

from table II with, again, the top line representing $y + x_N$, with the rightmost explicit digit being in the N th place. Since L and M are distinct, we have used two distinct differences from Table I (possibly different lengths of "2...2" in the third example), and thus two distinct additions from Table II. But this means that two top lines from Table II represent the same number $y + x_N$ in the same orientation. Examination of the various top lines shows this to be impossible. Thus the proof.

Notice that if $y = 41/81$, then $y + 13/81 = 2/3$ and $y + 13/27 = 80/81$, both Cantor numbers. So there exist translates which contain two Cantor numbers, but none with three or more.

This is the best possible in that differences between Cantor numbers comprise the entire closed interval $[-1, 1]$, and any bounded sequence has two elements within one unit of each other, so that some translate of that sequence will carry both of these elements into the Cantor set.

The related question, whether there exists a sequence each translate of which (through less than one unit, say, to avoid trivialities with boundedness) intersects the Cantor set in at least one point, can be answered negatively using measure theory, for such a construction would yield a covering of the unit interval by a countable number of copies of the measure-zero Cantor set.

Also solved by D. Borwein, J. W. Grossman, Nicholas Passell, Konrad Victor (Israel), and the proposer.

Completely Monotonic Functions

5852 [1972, 400]. *Proposed by C. H. Kimberling, University of Evansville*

Suppose f carrying $[0, \infty)$ onto $(0, 1]$ has alternating derivatives:

$$(-1)^k f^{(k)} \geq 0, \quad k = 0, 1, \dots$$

Prove $g(x) = (1 - f(x))/x$ has alternating derivatives on $(0, \infty)$.

I. Solution by Fred Schuurmann, Miami University, Ohio. Actually we may set $g(x) = (f(0) - f(x))/x$, where $f(x)$ maps $[0, \infty)$ onto $(0, f(0)]$. It is shown by induction that

$$xg^{(k)}(x) = -\{f^{(k)}(x) + kg^{(k-1)}(x)\}, \quad g^{(k)}(0) = -\frac{f^{(k+1)}(0)}{k+1}$$

$k = 1, 2, \dots$. Thus we have a first order linear differential equation in $g^{(k-1)}(x)$ and the solution is given by

$$g^{(k-1)}(x) = -\frac{1}{x^k} \int_0^x t^{k-1} f^{(k)}(t) dt.$$

Therefore g has alternating derivatives.

II. *Solution by O. P. Lossers, Technological University, Eindhoven, the Netherlands.* Since $f'(x) \leq 0$ and f is onto, $f(0) = 1$. By Taylor's formula

$$1 = f(0) = f(x - x) = \sum_{k=0}^n \frac{(-1)^k x^k}{k!} f^{(k)}(x) + \frac{(-1)^{n+1} x^{n+1}}{(n+1)!} f^{(n+1)}(\theta x),$$

where $0 < \theta < 1$. Hence

$$\begin{aligned} g^{(n)}(x) &= \frac{1}{x} \sum_{k=0}^n \binom{n}{k} (1 - f(x))^{(k)} (-1)^{n-k} x^{-(n-k)} (n-k)! \\ &= (-1)^{n+1} \frac{n!}{x^{n+1}} \left\{ \sum_{k=0}^n \frac{(-1)^k x^k}{k!} f^{(k)}(x) - 1 \right\} = -\frac{1}{n+1} f^{(n+1)}(\theta x). \end{aligned}$$

which proves the assertion.

Also solved by K. F. Andersen, Frederick Carty, Robert Heller, A. C. Hindmarsh, R. B. Kirk, Eitan Lapidot (Israel), B. E. Rhoades, Steven Russ, R. P. Soni, P. H. Young, and the proposer.

Order of Products of Elements in an Abelian Group

5853 [1972, 400]. *Proposed by Gomer Thomas, University of Washington*

Let x and y be elements of a finite Abelian group G with orders m and n respectively. Let q be the order of $\langle x \rangle \cap \langle y \rangle$, the intersection of the cyclic subgroups generated by x and y . Give the possibilities for the order of xy , in terms of m , n , and q .

Solution by the proposer. Let l be the least common multiple of m and n . If π is any collection of primes and h is any integer, let h_π be the (unique) integer defined by: (i) h_π divides h , (ii) every prime factor of h_π is in π , and (iii) no prime in π is a factor of h/h_π . Let σ be the set of primes which divide q but do not divide either l/m or l/n .

We shall prove that k can occur as the order of ab , $k = o(ab)$, if and only if k is of the form $k = l/s$, where (1) s divides q_σ , and (2) if $2 \mid q_\sigma$, then $2 \mid s$.

First, we show that $o(ab)$ must be of this form. Let $x = a^{m/q}$ and $y = b^{n/q}$. Both x and y are generators of $\langle a \rangle \cap \langle b \rangle$, so $y = x^r$, for some integer r satisfying $(r, q) = 1$. The smallest integer t such that $(ab)^t \in \langle a \rangle \cap \langle b \rangle$ is $t = l/q$, so $o(ab) = (l/q) \cdot o((ab)^{l/q})$. Now $(ab)^{l/q} = x^{l/m} y^{l/n} = x^{l/m+rl/n}$. Since $o(x) = q$, we have $o((ab)^{l/q}) = q/s$, where $s = (l/m + rl/n, q)$. Since $(l/m, l/n) = 1$ and $(q, r) = 1$, no prime dividing both q and one of l/m or l/n can divide $l/m + rl/n$. Hence s divides q_σ . If 2 divides q_σ , then $l/m, l/n$, and r are all odd, so 2 divides s .

Now let us show that any number of the specified form can occur as $o(ab)$. Specifically, let m, n, q, s be positive integers with m, n arbitrary, q a divisor of (m, n) , and s satisfying (1) and (2). Given $k = l/s$, let π be the set of primes dividing s , ρ be the set of primes dividing q_σ but not s , and τ be the set of primes dividing q but

not q_σ . Since $(l/n, q_\pi) = (l/n, q_\rho) = 1$, there exist integers t_1 and t_2 such that $t_1 l/n \equiv 1 \pmod{q_\pi}$ and $t_2 l/n \equiv 1 \pmod{q_\rho}$.

By the Chinese Remainder Theorem, there is an integer r which simultaneously satisfies the congruences:

$$r \equiv 1 \pmod{q_\tau}, \quad r \equiv t_1(s - l/m) \pmod{q_\pi}, \quad r \equiv t_2 l/m \pmod{q_\rho}.$$

It follows from these that $(r, q_\pi) = (r, q_\rho) = (r, q_\tau) = 1$, so $(r, q) = 1$. Moreover, the last two congruences give:

$$l/m + rl/n \equiv s \pmod{q_\pi} \quad \text{and} \quad l/m + rl/n \equiv 2l/m \pmod{q_\rho},$$

which implies that $(l/m + rl/n, q) = s$.

Now, let $G = \langle u \rangle \times \langle v \rangle$, where $o(u) = l$ and $o(v) = n/q$. Let $a = u^{l/m}$ and $b = u^{r/n}v$. Then $o(a) = m$, $o(b) = n$,

$$o(\langle a \rangle \cap \langle b \rangle) = q, \quad \text{and} \quad o(ab) = l/s.$$

Also solved by J. W. King, and by Brian Wesselink.

m-tuples and Their Branchings

5854 [1972, 523]. *Proposed by Stephen Gelbart, Princeton University*

Given a decreasing sequence of integers k_1, \dots, k_n , a *branching* is a sequence of integers k'_1, \dots, k'_{n+1} with $k_i \geq k'_i \geq k_{i+1}$. Upon successively branching $n - 1$ times one obtains a single integer; one calls a sequence of $n - 1$ successive branchings a *complete branching*. Show that there are

$$\prod_{1 \leq i < j \leq n} [(k_i - k_j + j - i)/(j - i)]$$

distinct complete branchings of a given sequence $\{k_i\}$.

Solution by Leonard Carlitz, Duke University. Let $N(k_1, \dots, k_n)$ denote the number of distinct complete branchings of the sequence $\{k_1, \dots, k_n\}$. Since

$$\prod_{1 \leq i < j \leq n} \frac{x_i - x_j}{i - j} = \left| \begin{pmatrix} x_j \\ i - j \end{pmatrix} \right| \quad (i, j = 1, 2, \dots, n),$$

the stated result is equivalent to

$$\begin{aligned} (1) \quad N(k_1, \dots, k_n) &= (-1)^{n(n-1)/2} \left| \begin{pmatrix} k_j - 1 \\ i - 1 \end{pmatrix} \right| \\ &= (-1)^{n(n-1)/2} \left| \begin{pmatrix} k_j - j - 1 \\ i - 1 \end{pmatrix} \right| \quad (i, j = 1, 2, \dots, n). \end{aligned}$$

For $n = 2$ we have

$$N(k_1, k_2) = \sum_{k_1 \geq k_1' \geq k_2} 1 = k_1 - k_2 + 1$$

in agreement with (1). Also it is evident from the definition that

$$(2) \quad N(k_1, \dots, k_n, k_{n+1}) = \sum N(k'_1, \dots, k'_n),$$

where the summation is over all k'_1, \dots, k'_n such that

$$(3) \quad k_1 \geq k'_1 \geq k_2 \geq \dots \geq k_n \geq k'_n \geq k_{n+1}.$$

Assume that (1) holds up to and including the value n . Then, by (2)

$$N(k_1, \dots, k_n, k_{n+1}) = (-1)^{n(n-1)/2} \sum \left| \binom{k'_j - j - 1}{i - 1} \right|,$$

where the summation is over all k'_1, \dots, k'_n satisfying (3). Hence

$$\begin{aligned} N(k_1, \dots, k_n, k_{n+1}) &= (-1)^{n(n-1)/2} \left| \binom{k_j - 1}{i} - \binom{k_{j+1} - j - 1}{i} \right| \\ &= (-1)^{n(n-1)/2} \left| \begin{array}{cccc} 1 & 0 & \dots & 0 \\ \binom{k_1 - 1}{1} & \binom{k_1 - 1}{1} - \binom{k_2 - 2}{1} & \dots & \binom{k_n - n}{1} - \binom{k_{n+1} - n - 1}{1} \\ \dots & \dots & \dots & \dots \\ \binom{k_1 - 1}{n} & \binom{k_1 - 1}{n} - \binom{k_2 - 2}{n} & \dots & \binom{k_n - n}{n} - \binom{k_{n+1} - n - 1}{n} \end{array} \right| \\ &= (-1)^{n(n-1)/2} \left| \begin{array}{cccc} 1 & -1 & \dots & -1 \\ \binom{k_1 - 1}{1} & -\binom{k_2 - 2}{1} & \dots & -\binom{k_{n+1} - n - 1}{1} \\ \dots & \dots & \dots & \dots \\ \binom{k_1 - 1}{n} & -\binom{k_2 - 2}{n} & \dots & -\binom{k_{n+1} - n - 1}{n} \end{array} \right| \\ &= (-1)^{n(n+1)/2} \left| \binom{k_j - j}{i - 1} \right| \quad (i, j = 1, 2, \dots, n+1). \end{aligned}$$

This completes the induction.

Also solved by Richard Stanley, and by the proposer.

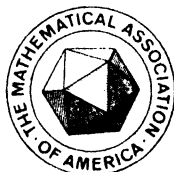
Editorial Note. Stanley derives the formula from Theorem 15.3 in his paper, *Theory and application of plane partitions*, Studies in Applied Math., 50 (1971), pp. 167–188, 259–279.

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF
THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 80



NUMBER 8

CODEN: AMMYAR

CONTENTS

On the Theory of Interest	DAVID GALE	853
The Nesting and Roosting Habits of the Laddered Parenthesis	R. K. GUY AND J. L. SELFRIDGE	868
Correction to "The Math. Societies and Associations in the U.K."	THOMAS WILLMORE	876
Squaring Rectangles and Squares	N. D. KAZARINOFF AND ROGER WEITZENKAMP	877
What Every College President should Know about Mathematics	J. G. KEMENY	889
Some Mathematical Verses		902
Remarks on Women in Mathematics	ASSOCIATION FOR WOMEN IN MATHEMATICS	903

MATHEMATICAL NOTES

A Note on Catalan Parentheses	JOHN RIORDAN	904
Determination of the Riemann Function	E. J. SCOTT	906
Inequalities for the Area of Two Triangles	L. CARLITZ	910
A Banach Space Characterization of the Space of Affine Continuous Functions on a Compact Convex Set	P. D. TAYLOR	911
On Lattice Points Inside Convex Bodies	A. M. ODLYZKO	915
Increasing Continuous Singular Functions	GERALD FREILICH	918

RESEARCH PROBLEMS

Questions on a Sequence of Ulam	BERNARDO RECAMÁN	919
Equitable Coloring	WALTER MEYER	920

CLASSROOM NOTES

The Lagrange Multiplier Rule	E. J. MCSHANE	922
The Mini-Max Property of the Tychonoff Product Topology	D. E. CAMERON	925
Another Approach to the Cubic Interpolating Spline	B. H. ROSMAN	927

(Continued on inside cover)

OCTOBER

1973

MATHEMATICAL EDUCATION

On Behavioral Objectives in Mathematics Education L. C. JANSSON AND R. T. HEIMER	930
Teaching a Computer-Oriented Laboratory Course for Ordinary Differential Equations H. E. WILLIAMS AND DELMER DE BOER	933
Multiple-Choice Examinations in Mathematics, not Valid for Everyone JERRY SILVER AND BERT WAITS	937
ELEMENTARY PROBLEMS AND SOLUTIONS	942
ADVANCED PROBLEMS AND SOLUTIONS	949
REVIEWS	953
NEWS AND NOTICES	969
MATHEMATICAL ASSOCIATION OF AMERICA	969
March Meeting of the Southeastern Section	969
March Meeting of the Southern California Section	971
April Meeting of the Iowa Section	972
April Meeting of the Nebraska Section	972
April Meeting of the North Central Section	973
April Meeting of the Ohio Section	973
April Meeting of the Oklahoma-Arkansas Section	974
May Meeting of the Illinois Section	975
Calendars of Future Meetings	976

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss.

Backlog: Main Articles 12 months, Math. Notes 13 months, Research Problems 7 months, Classroom Notes 11 months, Math. Education 10 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to ALEX ROSENBERG, Department of Mathematics, Cornell University, Ithaca, N.Y. 14850; NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D.C. 20036.

HARLEY FLANDERS, *Editor*

ALEX ROSENBERG, *Editor-Elect*

ASSOCIATE EDITORS

JOSHUA BARLAZ	J. G. HARVEY	SEYMOUR SCHUSTER
E. R. BERLEKAMP	ERIC S. LANGFORD	J. ARTHUR SEEBACH, Jr.
JANE W. DI PAOLA	P. D. LAX	E. P. STARKE
ROBERT GILMER	ARTHUR MATTUCK	LYNN A. STEEN
RICHARD GUY	M. W. POWNALL	JAMES WENDEL
RAOUL HAILPERN	GIAN-CARLO ROTA	

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are \$12.50. For nonmembers the subscription price is \$18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June-July, August-September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices.

Copyright © The Mathematical Association of America (Incorporated), 1973

PRINTED IN THE UNITED STATES OF AMERICA

ON THE THEORY OF INTEREST

DAVID GALE, University of California, Berkeley

1. Introduction. Every science is concerned with a certain set of phenomena. The “theory” is that part of the science which tries to explain why the given phenomena occur. Some scientific theories, even very profound ones like the theory of evolution, can be quite adequately presented in the language of ordinary discourse. But there are other theories which really cannot be understood except in the most superficial way without the use of mathematics. The man-in-the-street no matter how intelligent he is cannot be expected to get any real comprehension of, say, the theory of planetary motion, to say nothing of such things as quantum mechanics or relativity, without knowing some mathematics.

Economic theory, as practiced in the past by the great theorists (Marshall, Marx, Keynes, etc.) has been essentially a verbal theory. In teaching economic theory today there is an increasing tendency to “use” mathematics, but most of the basic ideas can still be conveyed by purely verbal arguments. An example is the classical theory of price formation. Since it will be needed later let me recall briefly how it goes. The price of any good is determined in such a way that the *demand* for it will be equal to its *supply*. One argues that if the good is priced too high people will refuse to buy it and sellers finding themselves with inventories rotting on the shelves will lower their prices. Similarly, if the price is too low sellers will find themselves constantly running out and will raise prices. This is really all one has to understand in order to grasp the essential notion of so-called *equilibrium* prices. (In comparatively recent times using quite advanced mathematics people have succeeded in proving that such equilibrium prices will exist under suitable conditions. However, it is clearly not necessary for a person to know the Brouwer fixed point theorem in order to comprehend the “law of supply and demand”.)

When one deals with economic phenomena which are a little more complicated than the simple one just described it may not be possible to come up with verbal man-in-the-street explanations. The next level of economic complexity, it seems to me, is what is called “the theory of interest” which, as we shall see, is really the theory of the behavior of prices over time. This theory attempts to explain why the institution of interest has been present in virtually all economic systems. This is actually a rather puzzling fact, for throughout history many of the “great thinkers” including Aristotle, Thomas Aquinas and, of course, Marx have condemned interest

David Gale received his Princeton Ph.D. under A. W. Tucker. He was at Brown University before joining the University of California, Berkeley. He held a Fulbright Research Scholarship at the University of Copenhagen, spent a year's leave at the RAND Corporation, held a Guggenheim Scholarship at Osaka University, and an NSF Senior Postdoctoral Scholarship at the University of Copenhagen. He is a Fellow of the Econometric Society. His research in mathematical economics includes the book *The Theory of Linear Economics Models*. Editor.

taking as unjustifiable and a form of extortion. The practice has often been attacked on moral and even religious grounds (Jesus drove the money lenders from the temple). Nevertheless according to Irving Fisher, author of a classical treatise on the subject, “despite all attempts to prohibit interest taking there is not and never has been in all recorded history any time or place without the existence of interest.”

Today most economists (but not all) agree that the explanation for interest is in fact the same as that for prices. It is necessary to have (positive) interest rates in order to balance supply and demand. Of course, demanding exorbitant interest is unjustifiable just as it would be unjustifiable to demand an exorbitant price for food or shelter, but—and this is probably less obvious—an interest rate which is too low may be just as harmful to an economy as one which is too high. The object of this essay is therefore two-fold, first to show, as convincingly as possible, why positive interest rates are required for the proper functioning of an economic system. This “qualitative” result is Theorem 2 of Section 5. We then examine a quantitative question. It turns out that the interest rate for an economy is what physical scientists call a “pure” or dimensionless number, meaning that its value is unchanged when we change the units in which other quantities are measured. This means it must be measuring something about the state of the economy. In Section 6 we describe exactly what this something is. Very roughly, an economy is shown to have the interest rate r if it is in a state where by sacrificing one unit of “consumption” in the present it can achieve a permanent increase of r units of consumption in every period thereafter.

I should perhaps say a word about why these results may be of interest to mathematicians. It seems to me there are two more or less opposite reasons for treating applications in a mathematics journal. The first is to show how “interesting mathematics” comes up in connection with some applied problem. There exists by now a fairly sizable literature of this kind where the *coup de grace* may involve an ingenious use of a fixed point theorem or some nonelementary measure theory. In the present exposition we are working the other side of the street. It is the problem which is of principal interest and one uses whatever mathematical methods seem appropriate for it, without regard to their aesthetic appeal. The analysis in the present case involves nothing more exotic than separating a pair of convex sets by a hyperplane at the critical moment. This is the type of mathematics that has dominated applications in economics throughout the “modern period” starting with Ville’s proof of the von Neumann minimax theorem. In any case the reader should not expect to see any mathematical rabbits pulled out of hats in the pages to follow and he should work through the succeeding analysis only if he is sufficiently interested in understanding what an interest rate is and why it occurs. (I should say that if one goes into the interest question in a more general setting, the mathematics picks up considerably. For example, the analysis in [1] and [3] is based on an elegant generalization of the Perron-Frobenius Theory of positive matrices.)

Finally there is the bibliographical question: whose theory is this anyway, and is it "correct?" In answer to the first question let me hasten to say, it's not mine. The subject of interest theory is possibly as old as economic theory itself, and I would hate to have the task of tracking down all the antecedents to the ideas to be presented here. In some ways, however, the present theory which has evolved almost entirely over the past twenty years is rather different from that of the classical writers. For example, one of the things the modern theory has made clear is the important relationship between the interest rate and the population growth rate (the former will in general exceed the latter as we shall see) a fact which the classical writers were apparently not aware of. As for the current literature my own education and involvement in these questions grew out of reading [1] and [2]. A book which is closely parallel to our formulation of the quantitative result is [4]. At least a dozen other papers over the past two decades have touched on one aspect or another of the present exposition. The particular package presented here however, has, as far as I know, not been given elsewhere.

As for correctness of the theory, questions in economics are never definitively settled the way they are in the physical sciences. People can always exhibit situations and circumstances under which any proposed theory breaks down, and I'm sure there are some economists who would dispute the ideas put forward here and claim that it is absurd to talk about interest in models which exclude such factors as say money or uncertainty about the future. My own feeling however, and I do believe many economists would concur, is that the theory to be described in the following sections gives about as convincing an explanation of the phenomenon under study as one can hope to get in an area as complicated as economics.

2. The laws of motion. Scientific theories often start with certain assumptions, its "laws of motion", and then proceed to deduce consequences of them. We shall follow this classical pattern and present here our basic economic laws, of which there are just two. Of course one cannot expect economic laws to hold exactly and invariably like, for example, the inverse square law of gravitational attraction, but if one is to accept the analysis which follows he must be willing to believe that, by and large, over sufficiently long periods of time these laws give a roughly accurate picture of economic reality. The first law has already been mentioned.

(I) *The supply and demand for each good in the economy are equal.*

We need not be concerned at present about the precise definition of supply and demand. Their meaning will be clear when we examine specific models. The idea is simple enough. If we think of the people in an economy as being producers and consumers we are requiring that the goods producers decide to produce shall be the same as the ones consumers want to consume. Again, the law is probably never satisfied exactly. On the other hand there is a sort of automatic feedback mechanism which assures that supply and demand will not remain too far apart, for if a good is

seen to be, say, oversupplied for any length of time the situation will be corrected either by producers supplying less of it or by lowering its price to increase demand. Please note that these considerations are completely independent of the type of economic system under consideration. Indeed, the laws of motion and in fact the whole theory to be presented here is “universal” in that it applies equally to a centrally planned socialist or *laissez-faire* capitalist economy, or anything in between.

The first law concerns only quantities of goods and would apply even to an economy without a price system. In order to state the second law we must assume that each good in the economy has a price and that goods are produced from other goods.

(II) *Among all possible methods of production, producers will choose those which maximize profits.*

At first glance this law may seem debatable. In a capitalist economy one might expect producers to maximize profits since presumably they will get to keep at least part of them, but why should the law hold in a planned economy in which profits revert to the state? The situation becomes more transparent if we rephrase the law to say that among all ways of producing a given commodity, producers should choose the one or ones whose cost is minimum. Stated in this way the law is hard to argue with, for consider the alternative. Why would authorities in a planned economy instruct producers to use a costly method of production when less costly ones were available? Presumably the planners control the prices as well as the choice of production technique and it hardly seems reasonable that these should be at “cross purposes” with each other.

3. Interest rates: The Equivalence Principle. What is an interest rate? Probably the most immediate way people think of interest is as a price one must pay for lending or borrowing money. Thus, to say that the interest rate is r percent per year means that if a person borrows (lends) d dollars he pays (receives) rd dollars per year throughout the duration of the loan. For our purposes it will be convenient to use a slightly different but equivalent definition based on what we call the *Equivalence Principle*.

Let us suppose there are n goods in the economy and the price of the i th good is p this year and will be p'_i next year.

EQUIVALENCE PRINCIPLE: *The following two economic worlds are equivalent:*

(A) *Money saved earns interest at the rate r per year.*

(B) *Money saved earns no interest but the price of the i th good next year is $p'_i/(1 + r)$ rather than p'_i .*

* Curiously, some recent criticism of (II) has been aimed at its relevance for a capitalist rather than socialist economy. J. K. Galbraith seems to be claiming that large firms deliberately use non-profit-maximizing techniques in order to increase their scale of operation. Other economists, however, consider this a rather dubious proposition.

In order for the reader to be convinced of this equivalence he need only think about it for a moment. Clearly the only thing a person who saves is concerned with is the *purchasing power* of his savings (we exclude consideration of misers). One sees at once that in either situation (A) or (B) the amounts of goods he can get for each dollar he saves is the same. Similarly, if he borrows a dollar to buy goods this year then in case (A) he repays $(1 + r)$ dollars next year while in case (B) he repays only the original dollar which, however, has now become worth $(1 + r)$ units of goods. But though the (A) and (B) worlds are equivalent, it is for many purposes much more convenient to adopt the (B) point of view. We illustrate this by a simple example which is standard in the theory of finance. For convenience we work with the *interest factor* $\rho = 1 + r$ rather than with r itself.

Suppose an individual anticipates a sequence of payments of money, his income, i_0, i_1, \dots, i_n over the coming $n + 1$ years. During that same period he plans to make a sequence of expenditures e_0, e_1, \dots, e_n . He knows he can borrow or lend money at the interest rate r and the question he asks himself is whether through a suitable program of borrowing and lending he can pay for the expenditure sequence (e_j) from his income sequence (i_j) . If the answer is yes, we call the expenditure sequence financially *feasible*. The direct approach to answering this question would go as follows: In year zero the planner gets i_0 and spends e_0 . His *savings* is therefore $s_0 = i_0 - e_0$, where of course s_0 can be negative, meaning that he must borrow in order to make his initial purchases. In either case, whether positive or negative, s_0 draws interest so his total wealth entering year 1 is $\rho s_0 + i_1$ of which he spends e_1 . Hence, his savings in year 1 is $s_1 = \rho s_0 + i_1 - e_1$, and so on. We thus have the system of equations

$$\begin{aligned} i_1 &= e_0 + s_0 \\ \rho s_0 + i_1 &= e_1 + s_1 \\ . &. \\ \rho s_{n-1} + i_n &= e_n + s_n. \end{aligned} \tag{3.1}$$

Now the condition that the expenditure sequence be feasible is by definition precisely that the number s_n be nonnegative, meaning that the planner is "solvent" at the end of the n year period. In order to see whether this will be so we multiply the j th equation of (1) by ρ^{-j} and sum, noticing that all the s_j terms, except s_n , cancel out, and we get

$$(3.2) \quad \sum_{i=0}^n i_j / \rho^j = \sum_{i=0}^n e_j / \rho^j + s_n / \rho^n,$$

and we see that s_n is nonnegative if and only if

$$(3.3) \quad \sum_{j=0}^n e_j / \rho^j \leq \sum_{j=0}^n i_j / \rho^j.$$

In economic terminology, the terms in the inequality are referred to as the *present*

value of the expenditure sequence and the income sequence respectively. We then have a very simple criterion for feasibility which says that the expenditure sequence is feasible if and only if its present value is less than that of the income sequence.

But now let us look at the same problem making use of the Equivalence Principle. There are now no interest payments. Instead all prices change from one year to the next by the factor $1/\rho$. If we think of the planner's income sequence as being his salary then since the price of his labor is falling at the rate $1/\rho$ his income sequence is $(i_0, i_1/\rho, \dots, i_n/\rho^n)$ and similarly the costs of the things he wants to purchase will be falling at the rate $1/\rho$, so the expenditure sequence becomes $(e_0, e_1/\rho, \dots, e_n/\rho^n)$. Now inequality (3.3) is simply the definition of feasibility, i.e., the requirement that expenditures shall not exceed income, and no calculations are necessary!

Henceforth we shall change over completely to the (B)-world formulation. The prices p'_i/ρ are sometimes called *present value* prices or *real* prices since they give the real value of next year's goods as compared with this year's. (The original numbers p'_i are called *money* prices.) We see now that a positive interest rate corresponds to falling real prices, and this allows us to phrase the interest rate problem in the following simple form. Why under normal conditions do real prices fall? (The distinction between real and money prices is, of course, crucial. As we are all too painfully aware, money prices have in recent years been rising, not falling, in most countries, but this "inflation rate" has been substantially below the "prime" rate of interest, so that the *real* interest rate has continued to be positive. Of course there have been historical cases of "runaway inflation" in which even real prices rise giving a negative real interest rate.) Our objective will be to show that if an economic system obeys the simple laws of the previous section then the real prices determined by these laws will in general have to decrease with time. Now if the economy itself is changing then one would expect prices to change too. The thing that is rather surprising, however, is the fact that even when the economy does not change in any way, the prices will. Specifically, we shall analyze economies in "steady states" in which quantities of goods produced and consumed remain the same period after period. When we calculate prices for this situation as determined by the two laws of the previous section it will turn out that the prices unlike the physical quantities do not remain constant but in general fall. This is the content of our first main theorem.

(Of course the economic world in which we live today is probably not even approximately in a steady state so one might argue that the theory presented here has no "relevance," to use the popular term. But anyone familiar with scientific method will realize that this kind of criticism is beside the point. It would be like criticizing Galileo's assertion that all bodies fall at the same rate in a vacuum on the grounds that on the earth's surface bodies never fall in a vacuum. The present exposition is not trying to explain interest in the economy of today or any other particular era but rather to isolate the basic universal causes of the interest phenomenon. For this purpose one studies not the real world but idealized models from which "irrelevant" disturbances have been deliberately eliminated. The steady state

in our problem serves the same purpose as the vacuum in the problem of falling bodies.)

4. Price theory for a simplified economy. We imagine an economy in which there are just three goods called *labor*, *capital*, and a third good on which people subsist called *consumption*. At each period of time t "nature" supplies a certain amount of labor L_t . For most of the analysis we shall assume that L_t changes from one period to the next by some constant *growth factor* γ so that the amount of labor at time t is $l\gamma^t$, where l is the fraction of *population* comprising the labor force.

Capital and consumption goods are not provided by nature but must be *produced*. The mechanism of production can be described in the following simple way. A productive *activity* \mathcal{A} is a 4-tuple of real members, written $\mathcal{A} = (-L, -K, C, K')$, where L and K are the *inputs* of labor and capital, and C and K' are the *outputs* of consumption and capital. The reason for the minus signs on the inputs will become apparent shortly. The ultimate goal of the economy is to produce consumption, capital being important only as an intermediate good which together with labor makes it possible to provide greater amounts of consumption. A *technology* \mathcal{T} consists of a set of activities, thus, a subset of 4-space. We make the following assumptions:

- (i) (homogeneity) if $\mathcal{A} \in \mathcal{T}$ then $\lambda\mathcal{A} \in \mathcal{T}$ for all $\lambda \geq 0$;
- (ii) (additivity) if $\mathcal{A}_1, \mathcal{A}_2 \in \mathcal{T}$ then $\mathcal{A}_1 + \mathcal{A}_2 \in \mathcal{T}$;
- (iii) (closure) \mathcal{T} is a closed set.

These assumptions mean, (i) an activity can be operated at any *level*, that is, multiplying inputs by some constant multiplies outputs by the same constant, and (ii) given two activities they can both be operated at the same time.

If \mathcal{A} and \mathcal{A}' are activities we say \mathcal{A}' *dominates* \mathcal{A} if $\mathcal{A}' \geq \mathcal{A}$ (the symbol $x \geq y$ means the vector $x - y$ is nonnegative and nonzero). Thus, one activity dominates another if from the same or smaller inputs it can produce the same or larger outputs. An activity in \mathcal{T} is called *efficient* if there is no other activity in \mathcal{T} which dominates it. If the objective of the economy is to maximize consumption output and minimize labor input it is clearly never desirable to use an inefficient activity. A concrete example may help to illustrate these ideas.

Example: We consider a technology in which there are two kinds of activities one for producing capital and the other for consumption. We suppose for simplicity that capital is produced by labor alone and, by suitable choice of units, we may suppose that one unit of labor is able to produce one unit of capital. This means that the capital producing activities of \mathcal{T} are of the form $(-L_K, 0, 0, L_K)$, where L_K is the amount of labor allocated to producing capital. Consumption is produced by capital and labor together in such a way that one unit of labor together with K units of capital produce K^α units of consumption where $0 < \alpha < 1$. In addition, capital used as input to production emerges from the productive process slightly *depreciated* so that there are only μK units of capital where $0 < \mu < 1$. Thus a typical consumption

producing activity is some multiple of an activity $\mathcal{A}_K = (-1, -K, K^\alpha, \mu K)$. Suppose now that at some instant of time there are L units of labor and K units of capital available and it has been decided to allocate L_c units of labor to producing consumption. Then in order to utilize all available capital one should operate activity $\mathcal{A}_{(K/L_c)}$ at level L_c giving the activity

$$(4.1) \quad \mathcal{A}_1 = (-L_c, -K, L_c^{1-\alpha} K^\alpha, \mu K).$$

Assuming full employment the rest of the labor force is used to produce new capital using the activity

$$(4.2) \quad \mathcal{A}_2 = (-(L - L_c), 0, 0, L - L_c)$$

and combining gives

$$(4.3) \quad \mathcal{A} = (-L, -K, L_c^{1-\alpha} K^\alpha, L - L_c + \mu K).$$

Since $K' = L - L_c + \mu K$ and $C = L_c^{1-\alpha} K^\alpha$ eliminating L_c gives

$$(4.4) \quad C = (L + \mu K - K')^{1-\alpha} K^\alpha$$

which holds for all (L, K, K') where $\mu K \leq K' \leq L + \mu K$. This describes the set of all efficient activities of the model.

It is clear from (4.4) that C is an increasing function of K and L and a decreasing function of K' as it should be. This leads to a second basic notion. The activity \mathcal{A} is called *unsaturated* if, roughly speaking, increasing any input enables one to produce more of all outputs. The precise definition is the following: let \mathcal{A}_i be the i th coordinate of \mathcal{A} . Then there exists $\tilde{\mathcal{A}}$ in \mathcal{T} such that $\tilde{\mathcal{A}}_i < \mathcal{A}_i$ and $\tilde{\mathcal{A}}_j > \mathcal{A}_j$ for $j \neq i$, for $i = 1, 2, 3, 4$. In the example it is clear that any efficient activity with positive coordinates is unsaturated.

We now introduce prices into the model and denote the price of a unit of labor by w (wage), of input capital by p , of output capital by p' and of consumption by q . A price vector π is then a 4-tuple $\pi = (w, p, q, p')$. It is crucial here that the price of input and output capital are not necessarily the same even though they are prices of the same physical good. The whole interest phenomenon, as we have seen, is concerned with prices which change with time and the point is that output capital becomes available only *after* input capital is applied. (It would clearly make no sense to have production take place instantaneously for then output capital could be fed back as input as soon as it was produced allowing infinitely large outputs in a single period.)

The *profit* of activity \mathcal{A} at prices π is simply the scalar product

$$(4.5) \quad \pi \mathcal{A} = -wL - pK + qC + p'K'.$$

The interpretation should be clear. The term $wL + pK$ is the cost of inputs and $qC + p'K'$ is the value of outputs and their difference is profit in its ordinary meaning. We now have a simple but basic result.

THEOREM 1. *If \mathcal{A} is efficient then there is a price vector $\pi \geq 0$ such that \mathcal{A} maximizes profits at prices π among all activities in \mathcal{T} . If in addition \mathcal{A} is unsaturated then π is positive.*

■ Let $S_{\mathcal{A}}$ be the set of all vectors x in 4-space such that $x \geq \mathcal{A}$. The definition of efficiency says precisely that $S_{\mathcal{A}}$ does not intersect \mathcal{T} . Further $S_{\mathcal{A}}$ is a convex set and so is \mathcal{T} from Assumptions (i) and (ii). It follows from the "Fundamental Theorem of Convexity" that there is a hyperplane H separating \mathcal{T} and $S_{\mathcal{A}}$. Letting π be the normal to H in the direction of $S_{\mathcal{A}}$ we have for any $\tilde{\mathcal{A}}$ in \mathcal{T} ,

$$(4.6) \quad \pi \tilde{\mathcal{A}} \leq \pi(\mathcal{A} + z) \text{ for any } z \geq 0$$

which implies $\pi \tilde{\mathcal{A}} \leq \pi \mathcal{A}$ and this is precisely the condition that \mathcal{A} is profit maximizing at the price π . Also $\pi \geq 0$, for let e_i be the i th unit vector. Taking z in (4.6) to be λe_i we have

$$(4.7) \quad \pi \tilde{\mathcal{A}} \leq \pi \mathcal{A} + \lambda \pi_i$$

for all $\lambda > 0$ so π_i cannot be negative.

To prove the last part of the Theorem, let $\tilde{\mathcal{A}}$ be the activity with $\tilde{\mathcal{A}}_i < \mathcal{A}_i$ and $\tilde{\mathcal{A}}_j > \mathcal{A}_j$ for $j \neq i$. Then

$$\pi \mathcal{A} = \sum_{j=1}^4 \pi_j \mathcal{A}_j \geq \sum_{j=1}^4 \pi_j \tilde{\mathcal{A}}_j$$

so

$$(4.8) \quad \pi(\mathcal{A}_i - \tilde{\mathcal{A}}_i) \geq \sum_{j \neq i} \pi_j(\tilde{\mathcal{A}}_j - \mathcal{A}_j).$$

but since $\pi \neq 0$ the right-hand side of (4.8) is positive, hence so is π_i . ■[†]

The reader should be aware of the economic significance of this rather simple theorem. If we accept the second law of motion that any activity used in the operation of our model should be profit maximizing then this fact alone determines what wages to pay to labor and what the return on capital shall be, at least for the case when the price vector π is unique (up to a multiplicative constant). As an illustration, let us consider the classical issue of a man who builds a machine and then hires labor to operate it in order to produce consumption. The question is then how much of the proceeds from production should go to the laborers and how much to the machine owner. Theorem 1 gives a simple answer. The workers get wL and the machine owner pK . Once again we emphasize that this result is "forced" on any economic system, capitalist or socialist, satisfying our second law of motion. Of course there are basic economic differences in the two systems when it comes to determining what happens to the quantity pK . In a pure capitalist economy this amount would go to increase the wealth of individual producers (or stock holders), while in the socialist economy

[†] Observe the notational break through. The initial symbol "■" does for the word "proof" the same job that the final "■" has been doing for "Q. E. D." all these years.

it would go to the state, since it is the machine owner, and the two types of economies will in many ways behave very differently—but not regarding the laws of price formation.*

Let us compute the prices in our special example. Given the efficient activity \mathcal{A} we must find prices π such that \mathcal{A} maximizes (4.5) subject to the equation (4.4). Substituting from (4.4) into (4.5) we have

$$(4.9) \quad \pi \mathcal{A} = -wL - pK + (\mu K + L - K')^{1-\alpha} K^\alpha + p'K',$$

where we normalize prices by taking $q = 1$. Setting partial derivatives of (4.9) with respect to L , K and K' equal to zero gives

$$(4.10) \quad \begin{aligned} w &= (1 - \alpha)(K/L_c)^\alpha \\ p &= [(1 - \alpha)\mu + \alpha/(K/L_c)](K/L_c)^\alpha \\ p' &= (1 - \alpha)(K/L_c)^\alpha, \end{aligned}$$

where we recall that $L_c = \mu K + L - K'$ is the amount of labor allocated to producing consumption. Note that the wage is equal to the price of output capital as it should be from our assumption that capital is produced by labor alone. The main quantity with which this paper is concerned is the *interest factor* $\rho = p/p'$ (see previous section). In the example, ρ is given by the expression

$$(4.11) \quad \rho = \mu + (\alpha/(1 - \alpha))(L_c/K).$$

We see that it is possible that the number ρ could be smaller than 1, for one can choose L_c/K as small as one wishes. This would correspond to a negative interest rate or equivalently a rising real price of capital. The following sections will show why this situation, though possible, is unlikely to occur.

Observe that the calculus technique used here will work whenever the set of efficient activities is *smooth*, that is, whenever there is a differentiable function ϕ such that the efficient activities are exactly those satisfying

$$(4.12) \quad \phi(L, K, C, K') = 0,$$

for note that prices must be such that $-wL - pK + qC + p'K'$ is a maximum subject to (4.12). Using a Lagrange multiplier λ in the standard way we have

$$(4.13) \quad w = -\lambda\phi_L, \quad p = -\lambda\phi_K, \quad q = \lambda\phi_C \text{ and } p' = \lambda\phi_{K'}.$$

(Actually we can choose $\lambda = 1$ since our technology is homogeneous so that if π is a price vector so is any positive multiple of it.) We shall return to this formulation in Section 6.

* A striking illustration of the “invariance” of interest with respect of the social system is the fact that savings accounts in mainland China pay a whopping 7% and there is no inflation so this is the real interest rate which is much higher than that of any capitalist country!

5. Steady states and the qualitative theory of interest. We now suppose that population in our model is growing by some fixed factor γ and that labor supplied in period t is $l\gamma^t$. The model is said to be in a *steady state* if it uses only the single activity \mathcal{A} in all periods, operating it at level γ^t in period t . We have not yet used the first law of motion and will do so now. The condition that supply equals demand in this model means that (A) the entire labor force is employed in all periods and (B) the amount of capital produced as output in period t is precisely the amount demanded as input in period $t + 1$. For a steady state $K_{t+1} = \gamma K_t$ and hence from (B) $K'_t = \gamma K_t$. Thus if we denote by $\mathcal{A} = (-l, -k, c, \gamma k)$ the activity used in period $t = 0$ then $\gamma^t \mathcal{A}$ is the activity used in period t . Note that the lower case letters l, k and c represent *per capita* quantities of labor, capital and consumption, and they remain constant in time.

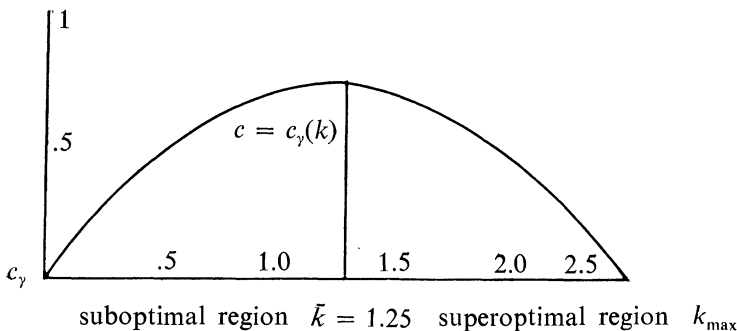
Now for each value k of per capita capital there will exist at most one efficient steady state corresponding to the activity with inputs l, k and output γk , having the largest possible value of c . We denote this value by $c_\gamma(k)$. The function $c_\gamma(k)$ is central to our analysis and it will be instructive to calculate it for the example of the previous section. For a steady state, equation (4.4) becomes

$$(5.1) \quad c_\gamma(k) = (l - (\gamma - \mu)k)^{1-\alpha} k^\alpha = l_c^{1-\alpha} k^\alpha,$$

where $l_c = l - (\gamma - \mu)k$. Notice that there are steady states for all k up to $k_{\max} = l/(\gamma - \mu)$ but not greater, for a *per capita* capital stock above this value could not be maintained even if all labor was allocated to producing new capital. Differentiating (5.1) gives

$$(5.2) \quad c'_\gamma(k) = (\alpha l/k - (\gamma - \mu))(k/l_c)^\alpha,$$

and we see that $c_\gamma(k)$ increases from 0 to a maximum as k goes from 0 to $\bar{k} = \alpha l/(\gamma - \mu)$ and then decreases back to 0 as k runs from \bar{k} to k_{\max} . We are assuming of course that $\gamma \geq 1$ and $\mu \leq 1$. (For the special case $\gamma = \mu = 1$, no population growth or depreciation, c_γ would be defined and increasing for all k .) Figure 1 gives the graph of $c_\gamma(k)$ for $\gamma = 1.03$, $\mu = .9$, $\alpha = \frac{1}{2}$, $l = 1$.



The steady state with $k = \bar{k}$ is called an *optimal* steady state. It represents the economic “millennium,” the “golden age” in which the standard of living has reached the highest possible sustainable value. In reality it seems unlikely that any society has ever achieved such a state and it is not even clear that it would be desirable to do so. As an example, imagine a society which has achieved a very high standard of living. Then a social planner points out that an even higher standard would be possible if all the present dwellings were torn down and replaced by even better ones, a process that might take several decades. Under these conditions the society might well decide that the slight improvement in future comfort was not worth the inconvenience of living in tents for twenty years.

From Figure 1 we also see that there are steady states with $k > \bar{k}$, but these are, from an economic point of view, highly unreasonable. They correspond to a society which has gone “beyond the millennium” and built up such a large stock of capital that it is wasting labor on maintaining the stock of capital instead of using it to produce consumption. In the extreme case, $k = k_{\max}$, all labor is allocated to maintaining the capital stock while the population starves to death!

The ideas illustrated above can be defined in general.

DEFINITION: A steady state \mathcal{A} is called **optimal** if it maximizes c_γ . It is called **sub (super) optimal** if there is another steady state $\tilde{\mathcal{A}}$ such that $\tilde{c}_\gamma > c_\gamma$ and $\tilde{k} > k$ ($\tilde{k} < k$).

As we have noted in connection with our example, optimal and especially super-optimal steady states, though conceivable, are economically unrealistic. The normal state of an economy is in the suboptimal region.

We can now give our fundamental qualitative result.

THEOREM 2. *If \mathcal{A} is an optimal steady state it has an interest factor equal to γ . If \mathcal{A} is sub (super) optimal then its interest factor will be strictly greater (less) than γ .*

Before proving this let us check it for our particular example. In a steady state we have seen that $L_c/K = l/k - (\gamma - \mu)$. Substituting this in (4.11) gives for the steady state interest factor

$$\rho = \mu + (\alpha/(1 - \alpha))(l/k - (\gamma - \mu))$$

and

$$(5.3) \quad \rho - \gamma = (\alpha l/k - (\gamma - \mu))/(1 - \alpha)$$

which is positive, zero, or negative according as k is less than, equal to, or greater than its optimal value $\bar{k} = \alpha l/(\gamma - \mu)$.

■ Let $\mathcal{A} = (-l, -k, c, \gamma k)$ be an efficient steady state. From Theorem 1 there exist prices $\pi = (w, p, q, p')$ such that

$$(5.4) \quad -wl - pk + qc + p'\gamma k \geq -wl - p\tilde{k} + q\tilde{c} + p'\gamma\tilde{k},$$

where $\mathcal{A} = (-l, -\tilde{k}, \tilde{c}, \gamma\tilde{k})$ is any other steady state. Recalling that $p/p' = \rho$ we get from (5.4)

$$(5.5) \quad q(\tilde{c} - c) \leq p'(\rho - \gamma)(\tilde{k} - k).$$

If \mathcal{A} is suboptimal then by definition we can choose \mathcal{A} so that $\tilde{c} > c$ and $\tilde{k} > k$ which implies $\rho > \gamma$. If \mathcal{A} is super optimal then we can choose \mathcal{A} so that $\tilde{c} > c$ and $\tilde{k} < k$ which implies $\rho < \gamma$.

Finally, suppose $\mathcal{A} = (-\bar{l}, -\bar{k}, \bar{c}, \gamma\bar{k})$ is optimal. For every $\mathcal{A} = (-l, -k, c, k')$ in \mathcal{T} consider the 3-vector $\beta = (-l, k' - \gamma k, c)$ and let \mathcal{B} be the set of all such β . Note that \mathcal{B} is convex and $\bar{\beta} = (-\bar{l}, 0, \bar{c})$ is in \mathcal{B} and by definition of optimality there is no β in \mathcal{B} with $\beta \geq \bar{\beta}$. It follows by the Separating Hyperplane Theorem as in Theorem 1 that there is a vector $\eta = (w, p', q)$ such that $\eta\bar{\beta} \geq \eta\beta$ for all β in \mathcal{B} . But this means

$$(5.6) \quad -w\bar{l} - p'\gamma\bar{k} + q\bar{c} + p'\bar{k} \geq -wl - p'\gamma k + qc + pk'$$

for all $(-l, -k, c, k')$ in \mathcal{T} , so letting $p = \gamma p'$ we see that \mathcal{A} has interest factor γ . ■

The theorem explains at least for steady states why we should expect interest rates to exceed growth rates, since, as we have seen, an economy will normally be in the suboptimal region. For an economy which is not in a steady state we do not even have in our model a sensible notion of interest rate since the prices of labor, capital and consumption may all be changing from period to period at different rates. Nevertheless, the general idea that (real) prices fall with time (though at different rates for different goods) still makes sense, and Theorem 2 gives a qualitative explanation of why this should occur for an economy in the suboptimal region where a higher future rate of consumption can be achieved only by first increasing the stock of capital.

6. What does the interest rate measure? Having learned why interest rates are positive it is natural to go further and ask what determines their magnitude. What distinguishes an economy where the interest rate is high from one where it is low? The qualitative analysis of the preceding section suggests various possibilities. We have seen for example that $\rho - \gamma$ is positive for $k < \bar{k}$ and negative for $k > \bar{k}$. It is natural to conjecture from this that ρ is a decreasing function of k and the magnitude of $\rho - \gamma$ measures in some way how close the economy is to its golden age. Notice that this is exactly the situation in our example as illustrated by equation (5.3). It has come as a rather recent shock in economic theory, however, that this plausible conjecture is not true. Even for our simple model one can concoct technologies in which the interest rate does not fall monotonically as a function of the level of steady state capital; so we must look elsewhere for the answer to our question.

Now there is one further qualitative property of steady states which can be proved. Namely the shape of the graph of $c_\gamma(k)$ of Figure 1 turns out to be typical.

THEOREM 3. *The function $c_\gamma(k)$ is strictly increasing (decreasing) if k is sub (super) optimal.*

■ This follows at once from (5.5), for from Theorem 2, if k is suboptimal then $\rho - \gamma > 0$, so if $\tilde{k} < k$ then $\tilde{c} < c$, and if k is super-optimal then $\rho - \gamma < 0$, so if $\tilde{k} > k$ then $\tilde{c} < c$. ■

Now suppose at some time t we have on hand a suboptimal stock of per capita capital stock k . From Theorem 3 it follows that by increasing capital to $k + \Delta k$ we can increase c_γ by some positive amount Δc_γ . The limit $\Delta c_\gamma / \Delta k = dc_\gamma / dk$, if it exists, is the slope of the graph of $c_\gamma(k)$ and measures the rate of increase of steady state consumption per unit increase in capital. Of course dc_γ / dk cannot be the interest factor for it is not a pure number but depends on the units used to measure c and k . But now let us carry the analysis one step further. In order to obtain an increase in input capital k_{t+1} next period it is necessary to increase output capital k'_t this period. Since $\gamma^t k'_t = \gamma^{t+1} k_{t+1}$, we have $k'_t = \gamma k_{t+1}$ so an increase of Δk_{t+1} next period requires an increase of $\Delta k'_t / \gamma$ this period. Finally, an increase of $\Delta k'_t$ this period requires that people sacrifice some consumption Δc_t , that is, they must accept Δc_t less than the steady state consumption $c_\gamma(k_t)$. The key question is then; how much of an increase in steady state consumption c_γ in the future can be obtained per unit sacrifice in consumption c_t today? That is, we wish to measure $\Delta c_\gamma / \Delta c_t$ or rather its limiting value $-dc_\gamma / dc_t$. (The minus sign occurs here because Δc_t was defined as a *sacrifice* of consumption.) We shall state the theorem relating this quantity to the interest rate for the special case in which the technology is smooth in the sense described in the previous section (a somewhat more complicated result holds in the nonsmooth case because of the possibility that the interest factor is not unique).

THEOREM 4. *For a smooth technology the quantity $(-dc_\gamma / dc_t)$ exists and is equal to $\tilde{r} = (\rho / \gamma - 1)$.*

For the special case when $\gamma = 1$ we have $\tilde{r} = r$, the ordinary interest rate. The content of the theorem can be described in economic terms in the following way. Given an economy in a steady state, it is possible instead of continually consuming all of c_γ to sacrifice some amount Δc . This “investment” of Δc units will provide (in the limit) a “dividend” of $\tilde{r}\Delta c$ units of additional per capita consumption forever after. The theorem also makes sense in the super optimal case in which \tilde{r} is negative. There it says that we can “have our cake and eat it too.” We can get an additional Δc units of consumption today and still have (in the limit) $-\tilde{r}\Delta c$ extra units ever after. This odd situation is due to the fact once again that in the super optimal case we have been wasting resources maintaining too much capital and we can be better off by “eating up” some of this burdensome excess.

■ The proof is an exercise in implicit differentiation. We rewrite equation (4.10) as

$$(6.1) \quad \phi(l, k_t, c_t, k'_t) = 0.$$

Since $\phi_{k'} = p' > 0$ we have

$$(6.2) \quad \frac{dk'_t}{dc_t} = -\phi_c/\phi_{k'} = -q/p' \text{ from (4.11).}$$

Next, as we have already observed, $k'_t = \gamma k_{t+1}$ so $dk_{t+1}/dk'_t = 1/\gamma$. Finally we must compute dc_γ/dk . For this purpose (6.1) becomes

$$(6.3) \quad \phi(l, k, c_\gamma, \gamma k) = 0$$

and differentiating with respect to k gives

$$(6.4) \quad \begin{aligned} \phi_k + \gamma \phi_{k'} + \phi_c dc_\gamma/dk &= 0 \text{ or from (4.11)} \\ dc_\gamma/dk &= (p - \gamma p')/q = p'(\rho - \gamma)/q \end{aligned}$$

so

$$\begin{aligned} -dc_\gamma/dc_t &= -(dc_\gamma/dk_{t+1})(dk_{t+1}/dk'_t)(dk'_t/dc_t) \\ &= (p'(\rho - \gamma)/q)(1/\gamma)(q/p') = (\rho/\gamma - 1). \blacksquare \end{aligned}$$

7. Concluding remarks. The whole of the preceding analysis has been carried out for a fictitious three-good world. It is natural to ask what parts of the theory carry over to a world such as the one we live in, in which there are thousands of goods. First it is important to point out those things that do not carry over.

Our fictitious model contained a fictitious good which we called "capital," and much of the theory depended on the amount of this good, the size of the "capital stock." In particular a crucial point of the analysis was whether the capital stock was above or below its "optimal" value \bar{k} . In a multi-good economy one can still talk of the stock of capital, this being all the factories, mines, buildings, transportation facilities and so on in the model, but it no longer makes sense to talk of the "size" of the stock. That is, there is no way in general to compare two different capital stocks and decide which one is "bigger." There have been various attempts by economists to find a suitable measure of the stock of capital but none has really succeeded and I doubt if the problem has a sensible solution.

Fortunately one does not need the concept of the amount of capital in order to generalize the theory of interest as presented above. The thing which does carry over to multi-commodity models is the concept of suboptimal, optimal, and superoptimal steady states. The idea is very simple. A steady state is *optimal* if there is no other steady state in which all members of the economy are better off. It is *suboptimal* if there exists a better steady state but it cannot be reached without making at least some members of the economy worse off during the time required to reach it. It is *super optimal* if there is a better steady state which can be reached at no sacrifice at all to any member of the economy. With these definitions the analogue of Theorem 2 carries over verbatim. The mathematics involved is considerably more complicated

values of the expressions for a given k . In this case the position of the innermost pair of parentheses is arbitrary, since

$$(2^2)^2 = 2^{(2^2)}.$$

We complete the solution by showing that for $k \geq 3$, the 2^{k-3} remaining values are all distinct. To prove this, in the evaluation each successive operation is either a squaring or an exponentiation base 2. We give the value, v_i , of an expression, in terms of its *second order exponent*, e_i ,

$$v_i = 2^{(2^{e_i})}.$$

Since

$$(2^{2^{e_i}})^2 = 2^{2^{e_i} \times 2} = 2^{2^{e_i+1}},$$

each operation is given by $e_{i+1} = e_i + 1$ or by $e_{i+1} = 2^{e_i}$. If there is a coincidence of values between two different k -level expressions, suppose that level $k(>3)$ is the lowest at which such a coincidence occurs. Since the $(k-1)$ -level expressions which gave rise to the coincidence are distinct, the equal k -level expressions have their last operations distinct; one an addition, the other an exponentiation. Thus $e+1 = 2^f$ where e, f are the second order exponents at level $k-1$. We may write $e+1 = 2^g + h$, where $1 \leq h \leq 2^g$, so that the last h operations were additions. At level 3 the second order exponent is 2, so $h \leq k-3$. Also $f \geq k-2$, because the second order exponent increases by at least 1 for each level from 3 to $k-1$. So

$$k-3 \geq h = 2^f - 2^g \geq 2^{f-1} \geq 2^{k-3} > k-3$$

and we have a contradiction. Hence all values above level 3 are distinct. The same method shows that for $a > 2$ the 2^{k-2} expressions all have distinct values.

We next ask how many k -level expressions there are if the $k-2$ pairs of parentheses are not necessarily nested. For $k \geq 4$ this number is strictly less than c_k since

$$(a^a)^{(a^a)} \quad \text{and} \quad (a^{(a^a)})^a$$

are equal, both having second order exponent $a+1$. This shows that exponentiation is not completely non-associative. A further problem is to count the distinct values of the k -level expressions for a particular value of a . The answers will be the same for all a such that there is no coincidence of value. We shall see that if there is a coincidence of value between a k_1 -level expression and a k_2 -level expression, then a coincidence occurs at all levels from $k_1 + k_2$ upwards. We assume a chosen so that no such coincidence occurs. Such a choice is possible since only a countable number are excluded. An outline of a proof of this is given by Göbel and Nederpelt [3].

We again work with second order exponents; now

$$(a^{(a^{e_i})})^{a^{(a^{e_j})}} = a^{(a^{e_i})(a^{e_j})} = a^{a^{e_i+e_j}}$$

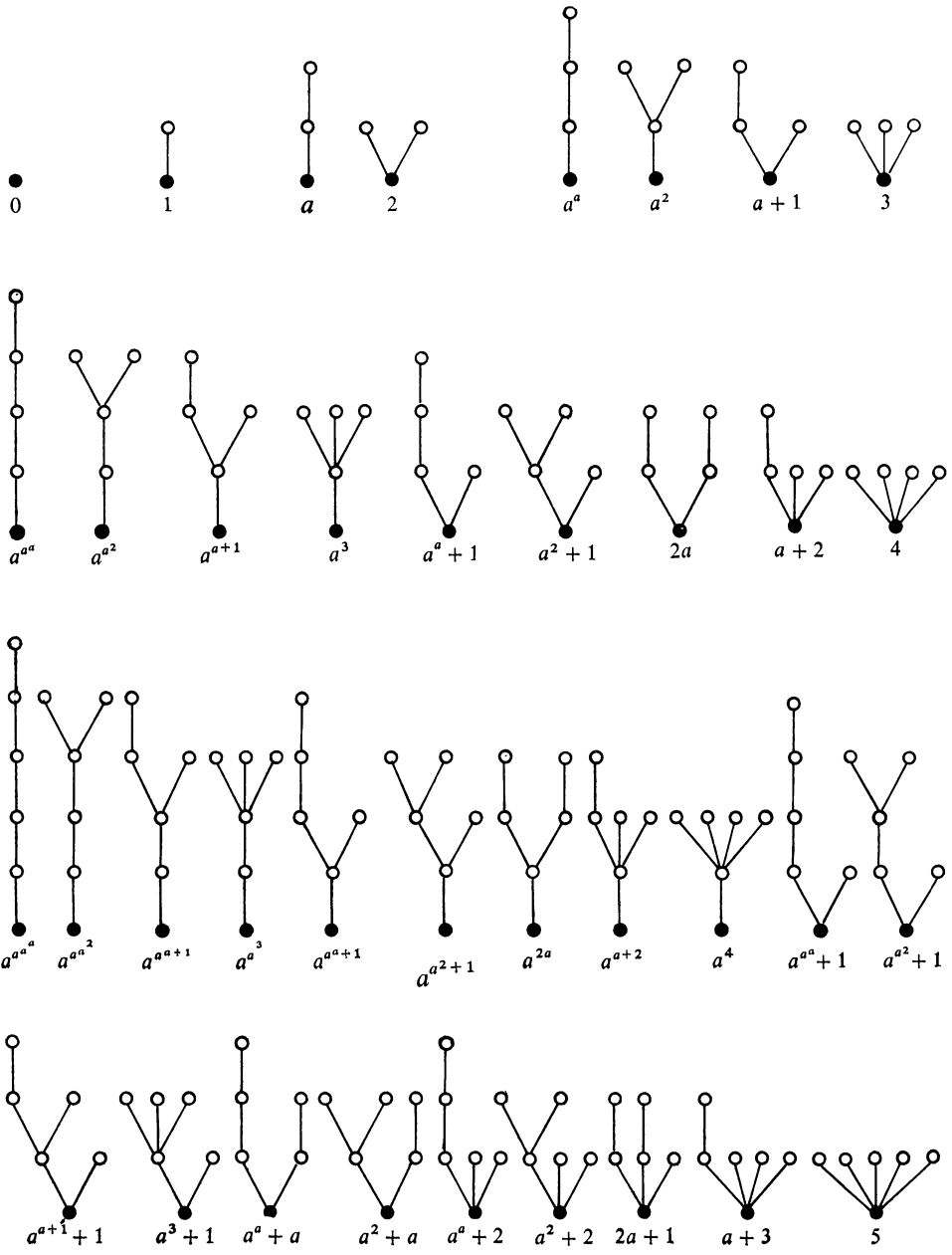


FIG. 1

so the second order exponents for level k are found by elementwise addition, for each i , of the pair of sets $\{e_i\}$, $\{a^{e_j}\}$ where i takes the values $1(1)k-1$, $i+j=k$ and e_i is a typical second order exponent for level i .

The sets $\{e_k\}$ for $k = 1(1)6$ are:

k	$\{e_k\}$
1	0
2	1
3	$a; 2$
4	$a^a, a^2; a+1; 3$
5	$a^{a^a}, a^{a^2}, a^{a+1}, a^3; a^a+1, a^2+1; 2a; a+2; 4$
6	$a^{a^{a^a}}, a^{a^{a^2}}, a^{a^{a+1}}, a^{a^3}, a^{a^a+1}, a^{a^2+1}, a^{2a}, a^{a+2}, a^4; a^{a^a}+1, a^{a^2}+1, a^{a+1}+1, a^3+1;$ $a^a+a, a^2+a; a^a+2, a^2+2; 2a+1; a+3; 5.$

On comparing the sequence of cardinalities of these sets with a prepublication version of N. J. A. Sloane's handy table [7], we learned what we should have guessed, that anything which nests is often associated with trees. In fact $|\{e_k\}| = r_k$, the number of non-isomorphic rooted, but otherwise unlabelled trees with k vertices. Knowing this, it is not difficult to see the correspondence between such trees and the sets as they are generated above. Exponentiation base a corresponds to growth, planting or grafting; addition corresponds to branching. Figure 1 shows all rooted trees with k vertices, $k = 1(1)6$. The parentheses are all nested except where the second order exponents are $2a$ at level 5 and a^{2a} , a^a+a , a^2+a and $2a+1$ at level 6. Methods of enumerating rooted trees are well known [4, 5]. The numbers may be calculated from the recurrence formula

$$r_k = \sum_{\pi(k-1)} \prod_i \binom{r_i + m_i - 1}{m_i},$$

where r_k is the number of rooted trees with k vertices, the sum is taken over all partitions $\pi(k-1)$ of $k-1 = \sum i m_i$ into $m_i (\geq 0)$ parts of size $i (\geq 1)$, and the binomial coefficient is the number of ways that m_i rooted trees, each with i vertices, chosen from the r_i possibilities with repetitions allowed, can be attached by m_i edges to a root to form a rooted tree with k vertices. The numbers for $k = 1(1)12$ are:

k	1	2	3	4	5	6	7	8	9	10	11	12
r_k	1	1	2	4	9	20	48	115	286	719	1842	4766.

In [5] the table is extended to $k = 26$.

To find the number of distinct values of the r_k expressions, when a takes a particular numerical value, is a more complicated problem. In the trivial cases $a = 1$ (or -1), only the value 1 (or -1) occurs at each level. If as usual $0^0 = 1$, then for $a = 0$ the values are 0, 1 for $k = 1, 2$ and both 0 and 1 for $k \geq 3$. We defer consideration of $a = 2$, which initiated our discussion, since it exhibits a special feature. We deal with $a = 3$, which will also serve as a model for larger integer values.

For $k=1, \dots, 6$, the numerical values of the second order exponents, when $a=3$, are

k	$\{e_k\}$
1	0
2	1
3	3; 2
4	27, 9; 4; 3
5	$3^{27}, 3^9, 81, 27; 28, 10; 6; 5; 4;$
6	$3^{3^{27}}, 3^{3^9}, 3^{81}, 3^{27}, 3^{28}, 3^{10}, 729, 243, 81; 3^{27} + 1, 3^9 + 1, 82, 28; 30, 12; 29, 11; 7; 6; 5.$

The semi-colons in this table and in the earlier one separate the contributions from the various partitions of $k - 1$ in the formula for r_k . So far the 1, 1, 2, 4, 9, 20 values are distinct at any one level, but the value 3 occurs at both levels 3 and 4; 27 and 4 occur at levels 4 and 5; and 3^{27} , 81, 28, 6 and 5 occur at levels 5 and 6. Note that the corresponding trees are those marked a and 3; a^a , $a + 1$ and a^3 , 4; a^{a^a} , a^{a+1} , $a^a + 1, 2a, a + 2$ and a^{a^3} , a^4 , $a^3 + 1, a + 3, 5$. They each arise from replacing the (sub)tree a with 3 vertices by the (sub)tree 3 with 4 vertices. Coincidences in value at the same level will occur whenever we have a tree containing tree a and tree 3 as disjoint subtrees, which yields a different tree when these two subtrees are interchanged. More generally, for any integer $a \geq 3$, the first coincidence in value, and the unique one at that level, occurs at level $a + 4$, the trees being those in Figure 2

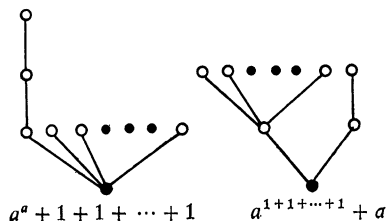


FIG. 2

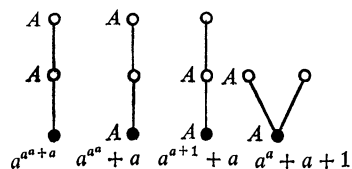


FIG. 3

They are obtained by grafting trees a and $1 + 1 + \dots + 1$ ($= a$), in either order, onto the two vertices of tree 1. To find all the coincidences at level $a + 5$ (i.e. level 8 if $a = 3$), we graft trees a and $1 + 1 + \dots + 1$ in every possible way onto two inequivalent vertices of each rooted tree with 3 vertices. Figure 3 exhibits the 4 ways

with pairs of vertices labelled A, A . At level $a + 6$ there are 16 coincidences, illustrated in Figure 4 and marked with the values of $a = 3$. More generally, the number of coincidences at level k would be the number of rooted trees with $k - a - 2$ vertices, with 2 inequivalent ones having indistinguishable labels. However, there are two further complications. The first is exhibited at level 10 for $a = 3$. If we start from the a -tree (Figure 5) and graft on $a, a, 3$ at its 3 vertices, we obtain three trees, each of value $3^{30} + 3$: there are 2 duplicates, where we would be counting 3. We must use the inclusion-exclusion principle and make allowance for the number of rooted trees with indistinguishable labels on 3 inequivalent vertices. For level 11 and $a = 3$ this amounts to 10 cases (Figure 6), the tenth arising from grafting $a, 3, 3$ onto the a -tree. The second complication is that new coincidences arise wherever a new power

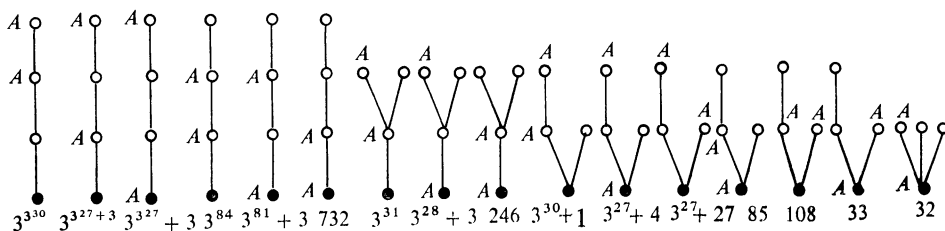


FIG. 4.

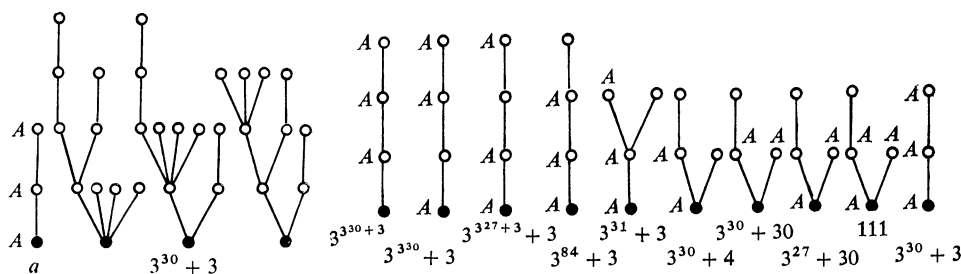


FIG. 5.

FIG. 6

of a occurs. For $a = 3$ this next happens at level 11 from the equality of a^2 at level 4 with $a + a + a$ at level 7 (Figure 7). Grafting these in either order onto the vertices of the 1-tree gives 2 non-isomorphic trees, each with 11 vertices and value $3^9 + 9$. More generally, this first occurs at level $2a + 5$.

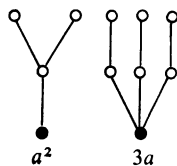


FIG. 7

For larger values of a , these events occur at correspondingly higher levels, so we are able to list the number of distinct values for $k = 1(1)11$ and $a \geq 3$.

k	1	2	3	4	5	6	7	8	9	10	11
$a = 3$	1	1	2	4	9	20	47	111	270	664	1659
$a = 4$	1	1	2	4	9	20	48	114	282	703	1787
$a = 5$	1	1	2	4	9	20	48	115	285	715	1826
$a = 6$	1	1	2	4	9	20	48	115	286	718	1838
$a = 7$	1	1	2	4	9	20	48	115	286	719	1841
r_k	1	1	2	4	9	20	48	115	286	719	1842

For $k \leq a + 3$, this number is the same as r_k . For $k = a + 4, a + 5, a + 6$, it is $r_k - 1, r_k - 4$ and $r_k - 16$. Thereafter the extra complications have to be taken into account. A more powerful enumeration could be made by an application of the Redfield-Pólya theorem, but technical difficulties will still arise.

We can answer the converse question: at what levels and with what frequencies does a particular value occur? Partition the value into parts which are powers of a ; similarly partition all exponents. Do this in every possible way. For example, if $a = 3$ then 28 can be expressed in 24 ways as

$$\begin{aligned}
 3^3 + 1 &= 3^{1+1+1} + 1 = 3^{1+1} + 3^{1+1} + 3^{1+1} + 1 \\
 &= 3^{1+1} + 3^{1+1} + 3 + 3 + 3 + 1 \\
 &= 3^{1+1} + 3^{1+1} + 3 + 3 + 1 + 1 + 1 + 1 = \dots
 \end{aligned}$$

so that 28 occurs (as a second order exponent) just at levels 5, 6, 11, 14–17, 17–23 and 20–29, i.e., it is duplicated at levels 17 and 20 through 23.

Finally we consider $a = 2$. Here there is an immediate coincidence at level 3, as we noted at the outset. In Figure 1, the a -tree and the 2-tree, have the same value. So we eliminate the former, and 'prune' all rooted trees, in the sense that wherever the a -tree appears, we replace it by the 2-tree. Such trees were called 'trimmed' by Göbel and Nederpelt [3]. As they pointed out, pruned trees can be enumerated by the same recurrence as for r_k , except that as we have replaced all a -trees by $(1 + 1)$ -trees, we have no contribution to any partition which contains a part of size 2. The corresponding numbers, s_k , of pruned trees with k vertices, are:

k	1	2	3	4	5	6	7	8	9	10	11	12	13
s_k	1	(1)	1	2	4	8	17	36	79	175	395	899	2074

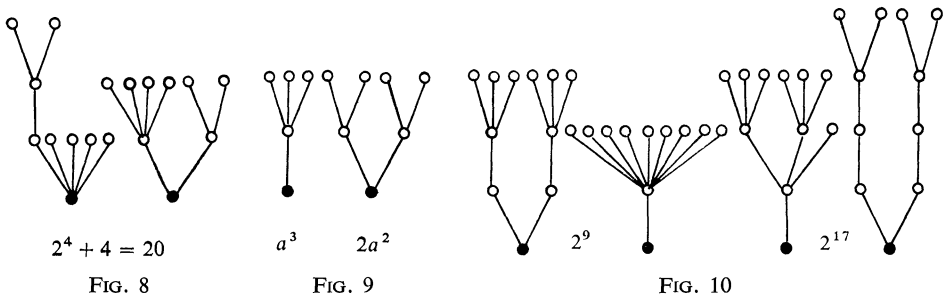
The parentheses mean that s_2 should be taken as zero in applying the recurrence relation.

For $a = 2$, the first few values of the second order exponents are:

k	$\{e_k\}$
1	0
2	1
3	2
4	4; 3
5	16, 8; 5; 4
6	65536, 256, 32, 16; 17, 9; 6; 4
7	265536, 2256, 232, 65536, 131072, 512, 64, 32; 65537, 257, 33, 17; 18, 10; 8; 7; 6.

The first coincidence is 4, at levels 4 and 5, so the first coincidence at the same level (above level 3) is $2^4 + 4 = 20$ at level 9 (see Figure 8). Complications of the first kind occur first at level $4 + 4 + 5 = 13$, and of the second kind at level $5 + 7 = 12$ from $a^3 = 2a^2$ (Figure 9). Note that in using Figure 4 to count duplicates at level 11 we ignore the 6th and 15th trees, since even after grafting they would contain an a -tree. But at this level there are two duplicates of the second kind, since (see Figure 10),

$$a^{2a^3+1} = 2a^{a^{a^2}} \text{ and } a^9 = 2a^{a^3}.$$



This gives the following numbers of distinct values of k -level expressions with $a = 2$.

k	1	2	3	4	5	6	7	8	9	10	11
$a = 2$	1	1	1	2	4	8	17	36	78	171	379.

There seems to be no simple characterization of what we might call *exponential numbers*, which lead to coincidences of value of k -level expressions. The coincidence may be between different levels in the first instance, but this will induce coincidences at the same level for all sufficiently large k , and the number of distinct values will be less than r_k for such k . The exponential numbers include all algebraic numbers, but do not form a field.

We list the numbers of distinct values of k -level expressions for the algebraic numbers $\frac{1}{2}(1 + \sqrt{5})$ and $\sqrt{2}$ and for the transcendental positive root of $a^a = 2$.

k	1	2	3	4	5	6	7	8	9
$a^2 = a + 1$	1	1	2	3	7	15	35	81	195
$a^2 = 2$	1	1	2	4	8	17	38	89	208
$a^a = 2$	1	1	2	4	8	17	39	90	213

We wish to thank John Riordan and the referee for suggestions.

R. K. Guy’s research supported by dwindling grant A-4011 of the National Research Council of Canada.

References

1. E. Catalan, Note sur une équation aux différences finies, *J. Math. Pures Appl.*, (1) 3 (1838) 508–516.
2. G. Eldredge, Nesting habits of the laddered parenthesis, *Problem E1903*, this MONTHLY, 73, (1966) 666; M. Goldberg, incomplete solution, *ibid.*, 77 (1970) 525–526; E. F. Schmeichel, comment *ibid.* 78 (1971) 298; completion of solution, *ibid.* 79 (1972) 395–396.
3. F. Göbel and P. R. Nederpelt, The number of numerical outcomes of iterated powers, this MONTHLY, 78 (1971) 1097–1103.
4. F. Harary, *Graph Theory*, Addison-Wesley, Reading, Mass., 1969, 187–190.
5. J. Riordan, An introduction to combinatorial analysis, Wiley, New York, 1958, 125–139.
6. ———, A note on Catalan parentheses, this MONTHLY, 80 (1973) 904–906.
7. N. J. A. Sloane, *A Handbook of Integer Sequences*, Academic Press, New York, 1973.

CORRECTION TO “THE MATHEMATICAL SOCIETIES AND ASSOCIATIONS
IN THE UNITED KINGDOM”

THOMAS WILLMORE, University of Durham, England

In this MONTHLY 79 (1972) 985–989, I stated that reviews of new mathematical books appear in the *Journal of the London Mathematical Society*. This used to be the case, but the London Mathematical Society now produces a very good journal, the *Bulletin*, which contains interesting information, lengthy expository articles and also the book reviews which previously would have appeared in the *Journal*.

I omitted all reference to the Edinburgh Mathematical Society, a Mathematical Society of long standing, which, although primarily concerned with mathematical research, has also had considerable influence on mathematics teaching. This justly provoked criticism from its President, Professor W. D. Collins, who incidentally extends a warm invitation to all members of the Mathematical Association of America to attend meetings of the Edinburgh Mathematical Society if they are able to do so. At least one Englishman will no longer identify “England” and “United Kingdom” in the future!!!

SQUARING RECTANGLES AND SQUARES

N. D. KAZARINOFF, State University of New York at Buffalo, and
ROGER WEITZENKAMP, The University of Michigan

1. Introduction. A **squared rectangle** is a closed rectangular region subdivided into a finite number of square regions that intersect only at their boundaries. The **order** of a squared rectangle is the number of its component squares. A squaring (of a rectangle) is **perfect** if no two component squares are congruent; otherwise it is **imperfect**. A **simple squared rectangle** properly contains no squared rectangle of order more than one. All other squared rectangles are **compound**.

Perfect squared rectangles of low order are easy to find, once one knows how to generate them. Perfect squared squares of low order are exceedingly rare at best. The perfect squared square of least order known has order 24 and is compound. It was found by T. H. Willcocks in 1948 [23, 24]. In 1965, W. T. Tutte [21] reported in this MONTHLY that A. J. W. Duijvestijn [9] had shown no perfect squared squares of order less than 20 exist. But, in fact, Duijvestijn only resolved the problem of determining all *simple* squared rectangles of order less than 20. We have recently [12] proved that there does not exist a *compound* perfect squared square of order less than 22. Thus there exists no perfect squared square of order less than 20, and Tutte's generous restatement of Duijvestijn's result is true.

Study of squarings of rectangles involves some graph theory, topology, combinatorics, number theory, and computer programming, which makes it an attractive subject. In this article we introduce the reader to the theory of squared rectangles, and we give an account of both recent and past results. The prerequisites we require are an elementary knowledge of topology, of how to solve a system of simultaneous linear equations, and of Kirchhoff's Laws. For an exposition less technical than ours, we refer the reader to an article by Tutte [19].

Although the study of squaring rectangles is old (see Section 6 for an historical account) a mathematical theory for squared rectangles is much younger. In 1940

N. D. Kazarinoff did his University of Wisconsin Ph. D. under R. E. Langer. He held positions at Purdue University and the University of Michigan before joining SUNY at Buffalo as Chairman of the Mathematics Department, and now also Martin Professor of Mathematics. He spent a year leave at the University of Wisconsin, was an exchange professor at the Steklov Institute of Mathematics, Moscow in 1960–61, and again in the spring semester of 1965.

He served as managing editor of the Michigan Mathematical Journal, as the consulting editor of Mathematical Reviews, as Chairman of the MAA Putnam Examination Committee, and on numerous MAA and AMS Committees. He is an elected member of CBMS, and in 1968 he received an award for Distinguished Undergraduate Teaching from the University of Michigan. His main research is in differential equations, and he is the author of *Geometric Inequalities* (1961), *Analytic Inequalities* (1961), and *Ruler and the Round* (1970).

Roger Weitzenkamp did his undergraduate and master's degrees at the University of Nebraska. He is a graduate student at the University of Michigan, and his main interest is combinatorics.

Editor.

Brooks, Smith, Stone, and Tutte [6] constructed an elegant, and indeed, definitive theory of squared rectangles. They related squaring rectangles to determining current distributions in certain electrical networks (planar graphs). These networks are composed of wires of one ohm resistance, except for one wire that contains a battery whose potential produces the current distribution. The central results of Brooks, Smith, Stone, and Tutte are: (1) *there exists a one-to-one correspondence between squared rectangles and certain equivalence classes of planar graphs*, and (2) *each simple perfect squared rectangle corresponds to an electrical network of unit resistances and battery that is equivalent to a 3-connected planar graph*. (See Section 3 for definitions of the terms **3-connected** and **planar graph**.) Our article is based on the theory of Brooks *et al.*

Somewhat after 1940, Tutte found, and later published [20], an algorithm for creating a complete list of 3-connected, finite, planar graphs. Duijvestijn [9] used Tutte's algorithm to search with an electronic computer for simple perfect squared squares by inductively creating a first portion of the list of all 3-connected planar graphs, an induction beginning with the 3-connected planar graphs of six and eight edges (Fig. 1).

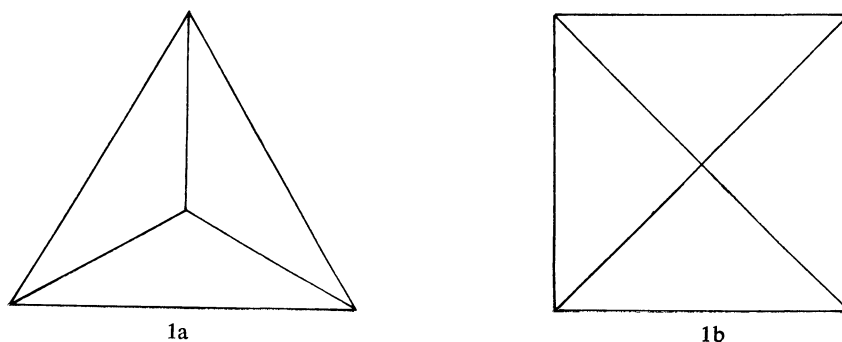


FIG. 1.

The theory of Brooks *et al.*, Tutte's theorem [20], and our extension of it (Theorems 2 and 4 below) provide an algorithm for generating all perfect squared rectangles—simple and compound. But almost nothing is known of the obvious problem: *given a closed rectangular region how can it be subdivided to yield a perfect squared rectangle?* Max Dehn [8] proved that *a rectangle can be squared if and only if its sides are commensurable*, and first R. Sprague [16] and, independently, Brooks *et al.* proved that *each rectangle with commensurable sides can be squared perfectly in infinitely many totally distinct ways*. But no one has found any algorithm for determining the perfect squaring of least order or even a reasonable estimate of that minimal order. For example, I. M. Yaglom [25] has shown that an a by b rectangle (a/b rational) *always can be subdivided to yield a perfect squared rectangle of order at most $13a^2b^2 - 11ab - 1$* , which for a 32 by 33 rectangle yields the upper bound

14,485,151. But in actuality the perfect squaring of a 32 by 33 rectangle of minimal order has order 9.

2. Generation of squared rectangles using Kirchhoff's Laws. The following example illustrates the method of Brooks *et al* for constructing squared rectangles from planar graphs. Tutte [19] has also written an elementary account of the relationship between planar graphs and squared rectangles. We consider the electrical network S^+ illustrated in Fig. 2.

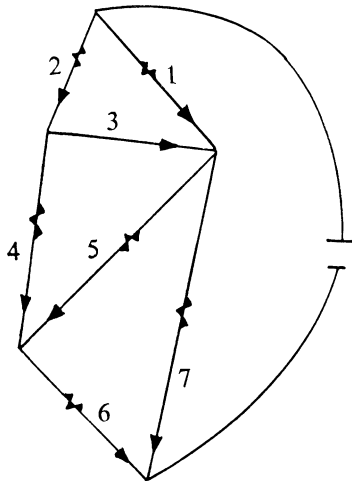


FIG. 2.

Suppose each resistance R_k is 1 ohm and that the current i_k in the resistance R_k is positive if it flows in the direction indicated and negative if it does not. Kirchhoff's Laws applied to a network are expressed as mesh equations (the change in potential around any closed path in the network is zero) and vertex equations (the flow of current into a vertex equals the flow out). For the illustrated network the vertex equations are:

$$i_1 + i_3 = i_5 + i_7$$

$$i_2 = i_3 + i_4$$

$$i_4 + i_5 = i_6.$$

The mesh equations are:

$$1 \cdot i_1 - 1 \cdot i_3 - 1 \cdot i_2 = 0$$

$$1 \cdot i_3 + 1 \cdot i_5 - 1 \cdot i_4 = 0$$

$$1 \cdot i_7 - 1 \cdot i_6 - 1 \cdot i_5 = 0.$$

These six equations are solvable for the seven currents (i_1, \dots, i_7) up to a constant factor of the unknowns, which we may choose so as to obtain a least solution in integers. This solution is (4,3,1,2,1,3,4). We now imagine that in the network S^+ each wire containing a resistance corresponds to a rectangle of width equal to the absolute value of current in the wire and height equal to the absolute value of the drop in potential over the wire. Since each R_k equals 1 ohm, the drop in potential numerically equals the current; and the associated rectangle is a square. The imperfect squared rectangle that is thus obtained from the network S^+ of Fig. 2 is illustrated in Fig. 3.

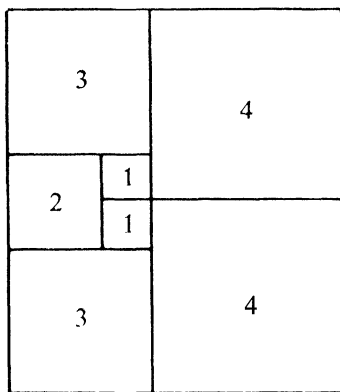


FIG. 3

Note that each horizontal line segment in this figure corresponds to a vertex of S^+ and that each square corresponds to an edge of S^+ . Given any planar electrical network of 1 ohm resistances and a battery, a squared rectangle can be derived from it in this way.

Let us vary this example. Suppose now $R_1 = b$ and $R_2 = R_3 = \dots = R_7 = 1$, that is, suppose that not all resistances in S^+ are 1 ohm. This adds one unknown, namely b , to the system of equations we obtain from S^+ and changes the first mesh equation to

$$b \cdot i_1 - 1 \cdot i_3 - 1 \cdot i_2 = 0.$$

To solve for (i_1, \dots, i_7, b) we need another equation. To provide it we add a "fixed ratio" condition, one that introduces the width-to-length ratio c of the "rectangled" rectangle to be derived from S^+ . This condition is:

$$c(b \cdot i_1 + 1 \cdot i_7) = i_1 + i_2.$$

For each c , the seven equations in the eight unknowns are again solvable up to a constant factor of the currents. For $c = 1$, a solution is (8, 4, 1, 3, 2, 5, 7, 5/8). Again we imagine that in S^+ each wire containing a resistance corresponds to a rectangle. Only this time one rectangle, the one corresponding to R_1 , is not a square. Because of the choice $c = 1$, the "rectangled" rectangle corresponding to S^+ is a square; see Fig. 4.

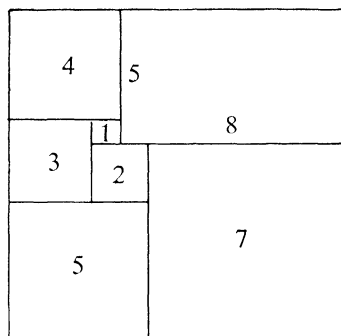


FIG. 4

The point of this variation of the example is that if we knew a squaring of a 5 by 8 rectangle, then it would yield a compound squaring of a square. We shall seek to make clear a process of exhaustively producing such compound squarings of squares.

There is, however, a difficulty arising from the first example. The resulting squared rectangle is compound, but the graph S^+ is 3-connected. We should like **simple** squared rectangles to correspond to 3-connected planar graphs, and **compound** squared rectangles to correspond to 2-connected planar graphs. A more careful scrutiny in the next section will allow us to make such a classification.

3. Graphs corresponding to perfect squared rectangles. A **finite planar graph** is a finite, planar collection of points called **vertices** and closed connected arcs called **edges**, together with a correspondence associating edges with vertices, namely, the vertices are the endpoints of the appropriate edges. A **loop** is an edge whose endpoints coincide. The **order** of a graph is the number of its edges. Throughout this article we deal only with connected, finite, planar graphs of positive order which we shall simply call **nets**, and we consider them as point sets.

A simple closed curve contained in a net that either contains no edges in its interior or all edges in the closure of its interior is a **mesh** of that net. If two vertices on the same mesh of a net are designated as poles, the net is a **polar net**. If A and B are polar nets, $A \subset B$, and A meets $B - A$ only at the poles of A , then A is a **polar subnet** of B . If S is a polar net with poles V and W , the **completion** S^+ of S is the net formed by joining the poles of S with one additional edge joining v and w .

Let S be a net. If there exists a vertex v of S such that $S - v$ is not connected, then S is **1-connected**. If S is not 1-connected, H and K are subsets of S , each containing at least two edges, and v and w are vertices of S , such that $S = H \cup K$ and $H \cap K = v \cup w$, then S is **2-connected**. If S is neither 1-connected nor 2-connected, then S is **3-connected**.

Our first theorem provides the means for defining a polar net corresponding to a squared rectangle. This theorem is a rewording of a theorem formulated by Tutte [17, §2.2] for triangles.

THEOREM 1. *For each squaring of a rectangle R with component squares S_j , there exists an orientation of the rectangle and a set of closed line segments p_i^σ ($\sigma = h, v$; $i = 1, 2, \dots, m_\sigma$), where m_h and m_v are positive integers, such that:*

- (a) *The union of the p_i^σ is the union of the sides of the component squares S_j , each side of each S_j being contained in some p_i^σ .*
- (b) *Each p_i^σ is horizontal or vertical as $\sigma = h$ or $\sigma = v$.*
- (c) *Two distinct segments have at most one point in common.*
- (d) *If w is a vertex of a square S_j but not a vertex of R , then w is an interior point of just one of the segments p_i^σ . If such a vertex w is common to four of the squares S_j , then w is an interior point of some p_i^v .*

Given a squared rectangle R , let $P = P(R)$ denote that polar net whose vertices correspond to the segments p_i^h and whose edges correspond to the squares S_j . If we consider P as an electrical network with unit resistance in each edge, a voltage applied to the poles of P induces currents in the edges which are proportional to the sides of the squares they represent. Brooks *et al* [6, p. 324] show that if R is simple, then P^+ is 3-connected, while Theorem 3 (below) shows that if R is compound, then P^+ is 2-connected. Since R can also be determined from P^+ , we have the desired correspondence between nets and rectangles which was mentioned at the end of the last section. Part (d) of the conclusion of Theorem 1 plays a key rôle in this correspondence.

DEFINITION. For $n \geq 5$, let \mathcal{L}_n denote the set of all finite planar graphs S such that:

- (a) S is a polar net of order n .
- (b) S^+ is 2-connected or 3-connected.
- (c) No two edges of S have the same pair of endpoints.
- (d) Each vertex of S that is not a pole is an endpoint of at least 3 edges.

It is easy to construct a family of nets which shows that \mathcal{L}_n is nonempty for each $n \geq 5$.

LEMMA. *Let R be a perfect squared rectangle of order n . Then $P = P(R) \in \mathcal{L}_n$.*

This lemma shows that the class \mathcal{L}_n was well chosen.

Proof of the Lemma. The net P is polar by construction, its poles corresponding to the p_i^h at the top and bottom of R . It has more than 5 edges because a perfect squared rectangle must contain at least nine component squares [6, p. 324]. To establish the second property of membership in \mathcal{L}_n it is sufficient to show that for any vertex v of P^+ there exists a circuit containing v and the poles of P . Such a circuit may be found by tracing a path from v to each pole via the corresponding squares in R , and including the edge $P^+ - P$. Finally, if either of the last two properties of membership in \mathcal{L}_n were violated by P , then two edges of P would carry the same (nonzero) current in the electrical model of P . This is not possible because R is perfect.

The set \mathcal{L}_n can be broken conveniently into four parts by the following theorem.

THEOREM 2. Each element S of \mathcal{L}_n satisfies exactly one of the following:

- (1) S^+ is 3-connected.
- (2) $S = X \cup x$, where $X \in \mathcal{L}_{n-1}$ and x is an edge added to a pole p of X in such a way that x connects p to one pole of S and $X \cap x = p$. The second pole of X is the second pole of S .
- (3) $S = Y \cup y$, where $Y \in \mathcal{L}_{n-1}$, the poles of Y are the poles of S , and y is an edge joining the poles of Y .
- (4) There exist integers m and k with $m, k \geq 5$ and polar nets A in \mathcal{L}_m and B in \mathcal{L}_k such that A^+ is 3-connected, and S^+ is formed by joining A and B at their poles.

We omit the proof of this theorem. It is long and somewhat tedious, involving counting and connectivity arguments.

The following theorem is implicit in [6, p. 323].

THEOREM 3. Let R be a compound perfect squared rectangle, and let $P = P(R)$. Then P^+ is 2-connected.

Proof. By the lemma, $P \in \mathcal{L}_n$ for some n . Therefore P^+ is 2-connected or 3-connected. Since R is compound, it properly contains a perfect squared subrectangle R_1 . Let $P_1 = P(R_1)$. From Theorem 1(d), we conclude that a vertex of P_1 that is not a pole of P_1 is incident only with edges corresponding to squares of R_1 . Thus P_1 is a polar subnet of P , and P^+ is 2-connected.

We shall call a net S in \mathcal{L}_n a T_i net ($i = 1, 2, 3, 4$) if S satisfies the i th conclusion of Theorem 2. To discover all compound perfect squarings of rectangles one need not consider T_1 nets because of the above theorem. Also, perfect rectangles corresponding to T_2 and T_3 nets consist of a square adjoined to one side of a smaller perfect squared rectangle, so that they are easy to find inductively. We are left with T_4 nets.

THEOREM 4. If Q is a compound perfect squared square of order n , then $P = P(Q)$ is a T_4 net of order n .

4. Gnomons. Defining a **gnomon** as the completion of a T_4 net, we know that to determine all compound perfect squared rectangles it is sufficient to create a hierarchal list of gnomons. Theorem 2 provides the means for doing this. Conclusion (4) of Theorem 2 describes the compound structure of gnomons, and all the conclusions describe the basic parts of gnomons. In this section we present an algorithm for creating a complete, hierarchal list of gnomons.

It is possible with forethought to eliminate certain portions of this list from consideration. For example if C is a polar net that corresponds to an imperfect squared rectangle, no gnomon G containing C can yield a perfect squared rectangle so long as the "battery" edge is an edge of $G - C$. After eliminating as many gnomons from the list as we easily could through mathematical analysis, we generated the remainder of the gnomons in the list having 22 or fewer edges by IBM-360 computer and dissected them one by one, also by computer. Over 17,000 gnomons

were dissected by the computer. The program we used to perform the dissections is a modification of Duijvestijn's program [9] that was written by James Reeds III.

We emphasize that when we dissected a gnomon of order n , we did so in all possible ways; that is, we solved the electrical networks determined by placing a battery in turn in each edge of the gnomon and unit resistances in each of the other $n - 1$ edges. The n squared rectangles resulting from these n dissections may be all different or there may be several alike. Each may be perfect or imperfect. The final result of this analysis of gnomons was the following theorem.

THEOREM 5. *There exists no compound perfect squared square of order 21 or less.*

We describe one method of constructing gnomons. Consider a T_4 net $S = S_0$ in \mathcal{L}_m , and let A, B, m , and k be as in conclusion (4) of Theorem 2. Since B belongs to \mathcal{L}_k , it is a T_i net for some i . If B is a T_2 or a T_3 net, remove the edge that corresponds to the edge x or y of Theorem 2 to obtain a net B_1 in \mathcal{L}_{k-1} . Repeat the procedure, if possible, to obtain $B_0 = B \in \mathcal{L}_k$, $B_1 \in \mathcal{L}_{k-1}$, \dots , $B_j \in \mathcal{L}_{k-j}$, \dots . The procedure terminates at some $S_1 = B_r$, where S_1 is either a T_1 or a T_4 net. Then the gnomon S^+ can be realized as the union of S_1 , a T_1 net $A_0 = A$, and $i_1 = l$ edges of the types of x and y in Theorem 2. If S_1 is a T_1 net, no further decompositions are needed. Otherwise, S_1 is a T_4 net, and S_1 is decomposed in the way S was decomposed. Repeat the procedure for each S_j ($j = 1, 2, \dots$) until a T_1 net, say S_r , is reached. At this final stage the original T_4 net S is described by:

- (1) a sequence of T_1 nets A_0, \dots, A_{r-1} , $A_r = S_r$,
- (2) a sequence of T_4 nets S_0, \dots, S_{r-1} , and
- (3) $q \equiv i_1 + \dots + i_r$ edges of the types of x and y in Theorem 2,

in such a way that for each $j < r$, the gnomon S_j^+ is the union of A_j , S_{j+1} , and i_{j+1} edges of the types of x and y ; see Fig. 5 for an example.

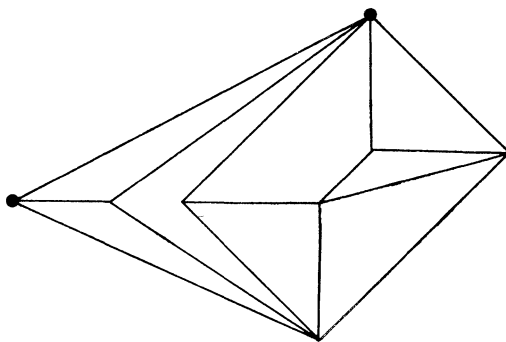


FIG. 5

By reversing the steps in the above procedure we reconstruct S^+ from the A_j 's and the extra edges. Indeed, beginning with arbitrary T_1 nets $\{A_j\}$ and as many extra edges as needed, we (theoretically) can construct all gnomons of a given order m .

Another procedure for searching for compound squared squares is illustrated by the second example in Section 2. One can substitute an unknown resistance in one or more wires in the electrical analogue of a net and solve that network by Kirchhoff's laws subject to the constraint that the resultant squared rectangle be a square. This method depends upon a knowledge of squared rectangles and, ultimately, simple perfect squared rectangles. In the example of Section 2 knowledge of a 5 by 8 rectangle is required. If there were one of order 17 or less, then it would yield a compound squared square of order less than 24, which could perhaps be perfect. (No such 5 by 8 rectangle does exist.) This method of search has been used by P. J. Federico [10].

5. 3-connected nets and simple squared rectangles. The fundamental data for generating compound squarings of rectangles are the 3-connected nets whose dissections yield simple squared rectangles. The basic theorem for generating 3-connected nets is Tutte's [20].

THEOREM 6. (Tutte). *Let G be a 3-connected planar graph with no loops, at least 4 vertices and such that no two edges have the same endpoints. Suppose further that G is not a wheel. Then either G or its dual graph can be derived from a simple 3-connected planar graph H by adjoining a new edge e to H whose ends are vertices of the same mesh of H and are not joined by an edge of H .*

Duijvestijn used this theorem to write a program for computer. Using the computer, he generated the 3-connected nets of orders less than 21 and showed there exists no simple perfect squared square of order less than 20.

We repeated some of his work. By computer we found and printed all dissections of rectangles corresponding to 3-connected nets with 17 or fewer edges. We also generated all such nets of orders 18 and 19, dissected them, and printed perfect dissections yielding ratios p/q with $p + q < 300$. (Duijvestijn [9] counts eight 3-connected nets of 12 edges. We found nine. All other counts agree.) We present some statistics arising from these data in Table I.

In Table II we give the Bouwkamp codes of the simple perfect squared rectangles of orders 16, 17 and 18 having sides with reduced ratios p/q such that $p + q < 30$. The Bouwkamp code of a 32 by 33 simple perfect squared rectangle of order 9, the least order possible, is: (18,15) (7,8) (14,4) (10,1) (9). The edge lengths of squares whose upper edges are segments of the same horizontal dissector are grouped in parentheses; the groups are listed in order of decreasing levels of the horizontal dissectors. These levels correspond to the levels of the potential in the corresponding electrical network.

In all perfect squared rectangles of small order (9 or 10 or 11) the largest subsquare appears in a corner. This phenomenon tends to persist, although occasionally the largest square appears at the "middle" of one of the sides. An example of a simple perfect squared rectangle of order 22 in which the largest subsquare appears in the "center" of the dissection is: (419,366,174,156,255) (18,138) (192) (39,216) (177) (53,505) (472) (393) (7,386) (133,379) (359,113) (246). This is a 1370 by 1250 rectangle.

TABLE I

No. of edges	Total No. of 3-connected nets	% of nets yielding at least one squared rectangle with sides having reduced ratio p/q such that					
		$p+q < 300$ (perfect)	$p+q < 30$ (perfect)	$p+q < 30$ (imperfect)	p or $q < 40$ (perfect)	$p/q=1$ (imperfect)	$p/q < 1/2$ (perfect)
10	2	50	0	100	50	0	0
11	2	100	0	100	0	0	0
12	9	22.2	0	66.7	0	11.1	0
13	11	54.5	0	63.6	9.1	9.1	0
14	37	32.4	0	48.6	0	10.8	2.7
15	79	22.8	0	35.4	0	3.8	3.8
16	249	14.1	0	20.5	1.2	4.0	6.8
17	671	14.0	.15	18.2	2.4	1.8	6.9
18	2182	11.5	.23	—	—	2.1	—
19	6692	8.7	.09	—	—	1	—
20	12,123*	5.5*	.1*	—	—	.8*	—
21	5,998*	3.0*	.1*	—	—	.5*	—
22	4,949*	2.2*	.04*	—	—	.3*	—

* only the number of nets sampled and percentages thereof

TABLE II

Order	p/q	Bouwkamp code
16	14/15	(87, 95) (39, 48) (40, 55) (27, 12) (3, 60, 25) (15) (10, 45) (42) (35)
16	11/18	(70, 73) (67, 3) (76) (39, 9, 7, 12) (2, 5) (11) (8, 85) (19) (58)
17	13/14	(51, 30, 88) (13, 17) (8, 5) (1, 16) (6) (56, 3) (9) (25) (19, 94) (75)
17	5/7	(17, 19, 27, 21) (15, 2) (13, 8) (5, 16) (1, 4) (33, 3) (7) (28) (23)
17	3/5	(40, 29) (13, 16) (36, 4) (2, 8, 3) (6) (19) (14) (12, 21) (39, 9) (30)
17	14/15	(145, 93) (45, 48) (42, 3) (23, 28) (110, 35) (18, 5) (33) (75,2) (20) (53)
17	11/15	(67, 52, 46) (6, 19, 21) (28, 30) (17, 2) (54, 13) (23) (41) (39, 8) (31)
18	9/10	(163, 98) (44, 54) (21, 23) (13, 41) (127, 57) (36) (8, 33) (19, 25) (70, 6) (64)
18	13/14	(123, 150) (91, 32) (6, 144) (23, 8, 1) (7) (15) (38) (80, 11) (9, 29) (20) (49)
18	14/15	(95, 61, 30, 54) (17, 13) (7, 6) (14, 3) (1, 5) (11) (59) (34, 52) (129) (111)
18	11/15	(76, 69, 119) (19, 50) (64, 12) (31) (21, 67, 112) (85) (39, 28) (11, 17) (135) (129)
18	11/15	(60, 29, 43) (15, 14) (1, 56) (16) (37, 39) (32, 5) (3, 11, 81) (8) (19) (51)
18	10/11	(129, 61, 70) (23, 29, 9) (20, 59) (17, 6) (11, 44) (28) (157) (5, 54) (49) (103)

An analogous phenomenon is a compound perfect squared rectangle composed of a perfect squared subrectangle surrounded by squares. Here is an example, first found by Federico (private communication). The elements of the squared subrectangle are set off by square brackets:

(390,389) ($[3 \cdot 14, 3 \cdot 18]$, 293) (292,98) $[3 \cdot 10, 3 \cdot 4]$ $[3 \cdot 7, 3 \cdot 15]$

$[3 \cdot 9, 3 \cdot 1]$ $[3 \cdot 8]$ (194), a 779 by 682 rectangle of order 15.

We have no statistics corresponding to those in Table I relative to compound squared rectangles because of the duplications that may occur among the various collections of gnomons of a given order that we separately constructed. We did observe, however, that compound perfect rectangles with $p + q < 30$ and $(p, q) = 1$ first occur with higher orders than in the case of simple perfect rectangles.

6. Historical notes. Max Dehn [8] proved in 1903 that a rectangle can be squared if and only if its sides are commensurable (see [25, §3] for a modern version of Dehn's proof). Yet no example of a perfect squaring was discovered until 1925 when Z. Moroń [14] found the 32 by 33 simple perfect rectangle of order 9. In 1930, M. Kraitchik [13, p. 272] quoted the famous N. Lusin to the effect that there exists no perfect squared square. Kraitchik only knew nontrivial examples of imperfect squared rectangles as well. Apparently Moroń's example was published in too obscure a journal (see also [7]). Finally, in 1939, the German geometer R. Sprague [15] discovered a compound perfect squared square of order 55 and side $5 \cdot 16 \cdot 29$, after he had attempted to prove Lusin's conjecture. Sprague built his example from two different 13 by 16 compound perfect squared rectangles and two squares. The next year Sprague [16] proved that each rectangle with commensurable sides has a perfect, perhaps compound, squaring, and, indeed, that each such rectangle has infinitely many totally distinct perfect squarings.

Almost simultaneously with Sprague and independently the paper by Brooks, Smith, Stone, and Tutte [6] appeared. Bouwkamp [1, 2] also found all the low order squarings of rectangles, but constructed no developed theory. Brooks *et al* did. They obtained Sprague's result, developed the analogy with electrical networks (which shows that each squared rectangle has commensurable sides and subsquares), and proved much more: there exists no perfect squared rectangle of order 8 or less; there exist two simple perfect squared rectangles of order 9; there exists a simple perfect squared square of order 55 and a compound perfect squared square of order 26—they gave examples [6, p. 333 and p. 334]. Using the electrical network analogy, C.J. Bouwkamp *et al* [3] gave a catalogue of all simple perfect squared rectangles of orders less than sixteen. In 1948 T. H. Willcocks, then a clerk for the Bank of England in Bristol, discovered a compound perfect squared square of order 24 and side 175 [18, 23, 24]. He holds the record still.

Duijvestijn [9] extended Bouwkamp's work. Federico [10, 11], Brooks [5], Bouwkamp [4], and John C. Wilson [21, 22] found interesting examples of squared squares and rectangles, mostly by computer. I. M. Yaglom [25] finally published the first book on squared rectangles. It contains much original material.

The outstanding open question remains: how to find the perfect squaring of least order of a given rectangle with commensurable sides. Perhaps it will prove easier to estimate closely this minimal order. The perfect squared square of least order will soon be found if computational difficulties are overcome or if much faster computers are built that will compute much more per dollar. It may well turn out to be Will-

cocks's gem: (64,56,55) (16,39) (38,18) (33,31) (3,4,9) (20,1) (5) (14) (30,81) (2,29) (35) (8,51) (43).

References

1. C. J. Bouwkamp, On the dissection of rectangles into squares, I-III, *Nederl. Akad. Wetensch. Proc.*, 49 (1946) 1176-1188; 50 (1947) 58-71 and 72-78.
2. ———, On the construction of simple perfect squared rectangles, *Nederl. Akad. Wetensch. Proc.*, 50 (1947) 1296-1299.
3. ———, A. J. W. Duijvestijn and P. Medema, Catalogue of simple squared rectangles of orders nine through fifteen, Department of Math. and Mech., Technische Hogeschool, Eindhoven 1960.
4. ———, On some special squared rectangles, *J. Combinatorial Theory*, 10 (1971) 206-211.
5. R. L. Brooks, A procedure for dissecting a rectangle into squares, and an example for the rectangle whose sides are in the ratio 2:1, *J. Combinatorial Theory*, 8 (1970) 232-243.
6. R. L. Brooks, C. A. B. Smith, A. H. Stone, and W. T. Tutte, The dissection of rectangles into squares, *Duke Math. J.*, 7 (1940) 312-340.
7. S. Chowla, The division of a rectangle into unequal squares, *Math. Student*, 7 (1939) 69-70.
8. Max Dehn, Über die Zerlegung von Rechtecken in Rechtecke, *Math. Ann.*, 57 (1903) 314-332.
9. A. J. W. Duijvestijn, Electronic computation of squared rectangles, Thesis, Technische Wetenschap aan de Tech. Hogeschool te Eindhoven, 1962.
10. P. J. Federico, Note on some low-order perfect squared squares, *Canad. J. Math.*, 15 (1963) 350-362.
11. ———, Some simple perfect 2×1 rectangles, *J. Combinatorial Theory*, 8 (1970) 244-246.
12. N. D. Kazarinoff and Roger Weitzenkamp, On existence of compound perfect squared squares of small order, *J. Combinatorial Theory*, B 14 (1973) 163-179.
13. Maurice Kraitchik, *La mathématique des jeux ou Récréations Mathématiques*, Stevens Frères, Bruxelles, 1930.
14. Z. Moron, O rozkladach prostokątów na kwadraty, *Przegląd. Matem. — Fizyczny*, 3 (1925) 152-153.
15. R. Sprague, Beispiel einer Zerlegung des Quadrats in lauter verschiedene Quadrate, *Math. Z.*, 45 (1939) 607-608.
16. ———, Über die Zerlegung von Rechtecken in lauter verschiedene Quadrate, *J. Reine Angew. Math.*, 182 (1940) 60-64.
17. W. T. Tutte, The dissection of equilateral triangles into equilateral triangles, *Proc. Cambridge Philos. Soc.*, 44 (1948) 463-482.
18. ———, Squaring the square, *Canad. J. Math.*, 2 (1950) 197-209.
19. ———, Squaring the square, "Second Scientific American Book of Mathematical Puzzles and Diversions," by Martin Gardner, Simon and Schuster, New York, 1961, 186-209. Reprinted from *Scientific American* November, 1958, 136-142.
20. ———, A theory of 3-connected graphs, *Indag. Math.*, 23 (1961) 441-455.
21. ———, The quest of the perfect square, this MONTHLY, No. 2, Part II, 72 (1965) 29-35.
22. ———, Squared rectangles, *Proc. I. B. M. Scientific Computing Symposium on Combinatorial Problems* (March, 1964) 3-9, I. B. M. Data Processing Div., White Plains, N. Y. 1966.
23. T. H. Willcocks, Problem 7795 and Solution, *Fairy Chess Review*, 7 (1948) 97, 106.
24. ———, A note on some perfect squared squares, *Canad. J. Math.*, 3 (1951) 304-308.
25. I. M. Yaglom, How to cut up a square? (Russian) *Nauka, Moskva* 1968.

WHAT EVERY COLLEGE PRESIDENT SHOULD KNOW ABOUT MATHEMATICS

JOHN G. KEMENY, Dartmouth College

(Invited Address, MAA Meetings, August, 1972)

Let me start with a very brief remark on the nature of the college presidency. The best characterization occurred to me somewhat accidentally when I was speaking to our California alumni during the primary races last spring. I found some remarkable similarities between the activities the political candidates were engaging in and what I was doing. Therefore, I should like to characterize the college presidency as the unique job in which you first get elected to office and then spend the rest of your time running for office.

It is a very peculiar multi-faceted job and it is not clear what kind of training is really good for a college president. But having a mathematician in such an office is a sufficiently rare event in the history of higher education that it might be worth considering whether there are ways in which being a mathematician is helpful to a college president. Obviously, there are a great many activities college presidents engage in where mathematics is totally irrelevant, but the opposite question is an interesting one. Therefore I will interpret the topic proposed for me by Professor J. L. Snell to mean "Are there certain things about mathematics that all college presidents should know?"

I shall briefly mention some trivia and then go on to a few interesting examples from my own experience.

Most people would say that it helps a college president to be able to read a financial statement. That is actually not as trivial as it sounds, because the purpose of most college financial statements is to meet a legal requirement, but to make sure that no one who reads the statement can find out what its contents mean. Secondly, it is good to be able to check the arithmetic in financial statements. I managed to find two errors in my first year as president, and I notice that this has had a salutary effect on the care with which these statements are prepared. But, more important is the fact that mathematicians are usually very good at explaining complicated mathematical things in fairly simple language to a non-mathematical audience.

John Kemeny was born in Budapest and did his undergraduate and graduate work at Princeton, where he received his Ph.D. He was in the US Army at the Los Alamos Project, and he served a year, as Albert Einstein's research assistant before becoming a Fine Instructor, and later Assistant Professor of Philosophy at Princeton. He has served Dartmouth College as Professor, Chairman of the Mathematics Department, Albert Bradley Professor, and President. He is a fellow of the American Academy of Arts and Sciences and has served the MAA on several committees, on the Board of Governors, and as Chairman of the New England Section.

He is the co-author or author of 13 books including the well-known *Introduction to Finite Mathematics*, *Finite Markov Chains*, *Denumerable Markov Chains*, and more recently *Man and the Computer: A New Symbiosis*. Editor.

That turns out to be a very useful asset to a college president, because you can take a highly complex statement and translate it for the faculty, students, and alumni.

I have a very strong feeling that decision-makers should not get their facts at second-hand. I practice that particular preaching, for example, by continuing to teach. I don't like to get feedback on what students are thinking by having a student tell his professor, who tells a dean, who tells a vice president, who tells the president. A great deal is lost in translation along the way. And I find that being able to remain active in the classroom is by far the best way of having your finger on the pulse of the campus. Similarly, there are many decision-making problems, and I will show you some examples, where being able to deal with the facts at first-hand gives you a much better feeling as to what the problem is and what the possible solutions are.

The most serious contribution that the mathematician-president can make is the fact that he knows something about model-building, and I would like to consider in detail some of the models that I have had to deal with in my first two and one-half years as a college president.

At the beginning of my term of office I realized that Dartmouth was facing a serious problem as far as the number of people on tenure was concerned. The size of the Dartmouth faculty was significantly increased in the 20's and the College was very good to its faculty during the depression and the war years, as a result of which in the early 50's the same faculty was still around and John Dickey, the previous President of Dartmouth, faced a horrendous re-building problem. Within a decade 80% of the permanent faculty would retire. He did a great job in that rebuilding, but it is almost impossible under those circumstances to avoid repeating past history. Once again you build up a strong young faculty, have them around for 30 more years, and face an impossible tenure situation 30 years later. Although the problem wasn't quite that extreme, it had to be tackled right away.

Let us consider a very simple model of the problem of ranks and tenure. A typical pattern of progression through the academic ranks is shown in Table 1. There is

TABLE 1

Instructor	??	Years
Asst. Prof. First Appointment	3	Years
Asst. Prof. Second Appointment	3	Years
Associate Prof.	6	Years
Professor	24	Years

a rather natural Δt for the model, namely 3 years. Furthermore, instructors are so ill-defined and mean so many different things that I am going to forget them in the simplified model. For the question of tenure, it's the non-tenure ranks versus the tenure ranks that are significant, so I'll lump the two tenure ranks together into one of 30 years average duration. The result is shown in Table 2. The actual

model we worked with was more sophisticated, but the essence of what we did can be brought out in this simplified model.

TABLE 2

API (Asst. Prof. First Appointment)	3 Years
APII (Asst. Prof. Second Appointment)	3 Years
Tenure	30 Years

Let us consider this model in some detail. What should the strategy for promotion and tenure be? You have three boxes into which you put not balls but faculty members. (I'm sorry—I've worked on probability problems too long.) The first one is the first appointment at assistant professor, the second one is second appointment at assistant professor, and the third is the tenure box with, say, x , y , and z people; and you have transition probabilities. With our three-year time cycle, and three-year appointments, everybody moves out of the first box in one time period. Say a fraction p_1 is reappointed, a fraction $1 - p_1$ leaves the institution, voluntarily or otherwise. In the second box a fraction p_2 is promoted, this time to tenure, and a fraction $1 - p_2$ is going to leave the institution. From tenure (since I have lumped together all the tenure ranks there is no promotion) some fraction p_3 will leave. This includes retirements, deaths, and going to another institution because of a better offer. The transition diagram is shown in Table 3.

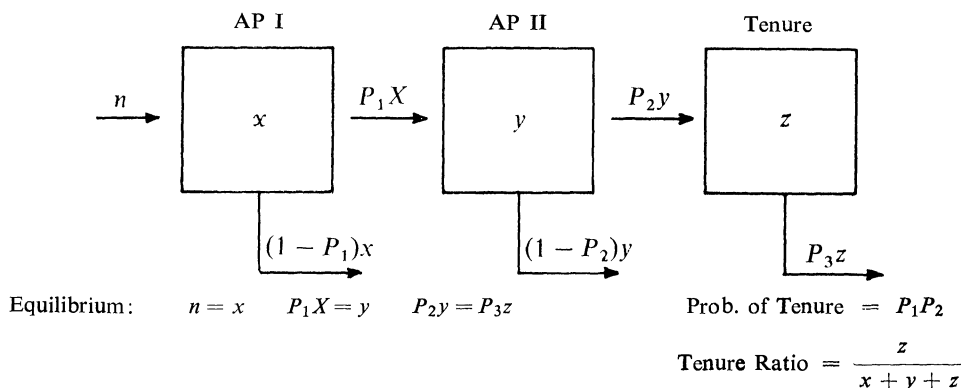
One can control p_1 and p_2 by institutional policy. One really has little control over p_3 . The problem is to relate the choice of parameters to the two most interesting quantities: the fraction of faculty on tenure and the probability that a new assistant professor will reach tenure.

The equilibrium conditions are easily found. For simplicity I'll assume that the total number of faculty members is fixed during this period, so the number coming in will have to equal the number going out. Let n new faculty members come in during 3 years, all as API. Since everybody leaves the first box and all the new ones come in at the beginning level, n must equal x . Next, from the first box p_1x goes to the second box, and everybody from the second box leaves at the end of three years, so p_1x must equal y . And since p_2y is the number entering the tenure ranks and p_3z is the number leaving, these two numbers must be equal. There is an additional condition that the number of new people coming in must exactly equal the numbers that go out, but that turns out to be a consequence of the other three equations and is therefore redundant.

What is the probability that someone coming in as a new assistant professor is going to reach tenure? With probability p_1 he will get promoted once, with probability p_2 he will get promoted twice and therefore, on the average, p_1p_2 is the fraction that is going to reach tenure from an initial appointment. Clearly, if you are going to care about your faculty members you'd like that quantity to be high. The tenure ratio is simply the number of tenure people divided by the total number, so it is

$z/(x + y + z)$. These results are summarized in Table 3.

TABLE 3 PROMOTION STRATEGY



Now the question is what can you do with the amount of freedom you have and what are the implications of various policy decisions? From the equilibrium equations we can express y and z in terms of x . Substituting in the formula for the tenure ratio and simplifying we find that

$$\text{Tenure ratio} = \frac{p_1 p_2}{p_3 + p_1 p_3 + p_1 p_2}.$$

Note that the probability of reaching tenure, $p_1 p_2$, occurs in two places in the formula.

Our aim is to make the quantity $p_1 p_2$ as high as possible, and to have the tenure ratio as low as possible. Suppose you have set the value of $p_1 p_2$ as high as you can, what are the remaining quantities? They are all in the denominator, and all with plus signs; so you would like to make them large in order that the tenure ratio be low. Now we'd love to fiddle around with p_3 but that's a dangerous business. Clearly the number of people who retire through mandatory retirement is fixed. In equilibrium, if we are talking about 30 years' tenure, in a three-year period roughly one-tenth will retire. In addition to that some fraction will leave voluntarily, say another 5%, so p_3 might be something like .15. But beyond that, all you can do is make life miserable for senior faculty members so that they will decide to leave the institution. Let us, therefore, assume that you have no control over p_3 . That still leaves you p_1 , and the first surprising conclusion is that p_1 should be as high as possible.

This calculation quite clearly shows there is every advantage to making p_1 as close to 1 as possible (for simplicity I'll use the value 1, though in practice that isn't likely). That is, a first term assistant professor should have a very easy time getting a reappointment. Obviously this is a very popular decision in a difficult job market. Setting $p_1 = 1$ will both help you in increasing the probability of reach-

ing tenure and help to decrease the tenure ratio. So here is a nice counter-intuitive result which actually led to a change of policy at Dartmouth College.

With $p_1 = 1$, the tenure ratio comes out to be $p_2/(2p_3 + p_2)$ and since you have very little freedom about p_3 , the ratio is determined by p_2 . Substituting a rough value .15 for p_3 , we obtain a tenure ratio of $p_2/(p_2 + .3)$. The ratio is monotone increasing in p_2 . And since $p_1 = 1$, the probability of reaching tenure is precisely p_2 . Now you have a conflict. On the one hand, to protect the future of the institution, you don't want the ratio to rise too high; on the other hand, you would like to make p_2 as high as possible, to make the institution attractive to new faculty members. I'll give you two typical values. If the tenure ratio is 50% the probability of reaching tenure is only 30% according to this model. On the other hand, if the tenure ratio goes up to 60%, then the probability of reaching tenure in equilibrium is 45%, which is reasonable for a new assistant professor at a good institution. While we did not have an equilibrium situation at Dartmouth and therefore our value is not as favorable, a calculation like this persuaded us to allow the tenure ratio to creep up over a decade to 60% so that we can give a decent chance for new assistant professors to stay here.

That's my first example of a model. The model is not terribly complicated; it does not use advanced mathematics; and yet it uses the kind of argument that someone not trained in mathematics is not likely to come up with or even be able to follow. Therefore I think it is an example that is legitimate under the topic "What mathematics should all college presidents know?"

Let me go to a second model, which I'll do more briefly. We are in a period of expansion in the Dartmouth faculty since we are about to go on year-round operation. That raised the question of how new faculty should be distributed amongst the existing departments. That's an old debate on almost every academic campus and any time you try to come up with an agreement as to what's an equitable formula, you have a losing fight. I know because I tried to get agreement and I lost the fight!

But it is still an interesting problem and therefore I asked myself a slightly different question: "Can one give a rational reconstruction for the way we are now assigning numbers of faculty members to departments?" One knows that the process is not totally rational. A great many conditions influence it; the persuasiveness of a departmental chairman, the prejudice of a dean and accidental conditions have an effect on the size of the departments. Nevertheless, can one look at the facts, identify those factors that could influence such a decision, and come up with a rational reconstruction as to how the assignments were made?

We tried coming up with a set of relevant factors. At first we identified too many factors, and if you have too many possible factors you can explain absolutely everything, including things that are just plain wrong in the system—you pick up a great many accidental correlations. In a way the only difficulty was to cut down the number of factors to a reasonable list which still gives you a good explanation (but not too

good). We came up with seven factors (see Table 4) in terms of which we got an amazingly good linear fit on the distribution of faculty members.

TABLE 4 — FACULTY LOAD FACTORS

1. Students in regular sections
2. Students in lecture sections
3. Students in labs
4. "Must be small" sections
5. Majors
6. Graduate students
7. Constant

The method of "best linear fit" assigns coefficients C_1, C_2, \dots, C_7 to the seven factors so that the linear combination "fits" the actual number of faculty members as closely as any linear formula can. The coefficients are the weights assigned to the various factors, and their interpretation is quite interesting. I shall discuss several coefficients, taking the liberty of rounding the answers.

The coefficient C_1 may be interpreted as assigning one full-time faculty member for each 150 students in "normal" sections. Thus a department "earns" a faculty member for teaching six sections of 25 students each or for five sections of 30 students. The value of C_2 is smaller; there must be 250 students in lecture sections before the department is given a full faculty member. Additional faculty is assigned for lab sections, for supervision of majors and of graduate students (C_3 , C_5 and C_6). Coefficients C_4 and C_7 deserve special mention.

You clearly don't want to reward a department simply for giving many small sections, on the other hand, there are departments that are forced to give small sections. For example, the Faculty at Dartmouth has voted that a Freshman Seminar cannot have more than 15 students in it, and clearly this must be taken into account in the loads of departments. A department is assigned one faculty member for every 12 sections that must be small. To interpret this we must recall that our formula is additive. For example, if a department has 150 students in 12 sections that must be small, it will be given one faculty member for the 150 students taught, and an additional faculty member for the 12 special sections. This seems to be fairly equitable recognition of the extra work involved in teaching many small sections.

Let us now look at the constant term $C_7 = 1$. That is, each department is given one faculty member irrespective of load. I think the interpretation is simply the overhead of having a department. So, if this reconstruction is at all reasonable, every time you create a new department, you commit something like one full-time faculty member just because it is a separate department, quite independently from any teaching load the department may carry.

This reconstruction (and let me remind you again that nobody actually makes the decisions this way) is a rational reconstruction of how we *could* explain the size of different departments. You could pick up a certain number of faculty members

because of a certain number of students, plus some additional ones because of students electing a certain special section, plus additional ones because of majors or graduate students, plus one for each department, etc. Although this reconstruction came out automatically from a computer run, the weights correspond reasonably well with our intuition as to how we might have done this assignment. Therefore there is something to it. And we found one other very convincing piece of evidence. The formula doesn't of course fit things perfectly, and there were three or four notable examples of departments having too many faculty members, or too few, and in every case where that happened, they were departments where the deans knew that the department was either overstaffed or understaffed.


To me it is quite surprising that you can do that well in a rational reconstruction, and we are planning to use this to monitor the growth of the Dartmouth Faculty in the next few years, as enrollment increases, and as there is the possibility of a shift in enrollments because of the presence of a significant number of women. (Though I suspect the shift will be less than most people predict—I think students take courses because of the reputation of departments on campus and not because they are men or women.) But whatever may happen, this model will give us a check as to whether we are allocating enough faculty members to departments that are being heavily hit by new enrollments.

So this is a second mathematical model, of a somewhat different type, that has a great deal to do with long-range planning at an institution. As a third example I picked a combinatorial problem arising from our decision to go on year-round operation. Why are we going on year-round operation, aside from the fact that, as you can see, Hanover is a very beautiful place in the summer? The goal is to accommodate, in addition to 3000 male undergraduate students, about 1000 women students and to do that without building any additional buildings—for the simple reason that we can't afford it. The idea was to spread out the academic year from three terms to four terms and accommodate a larger number of students in the same space (but with a larger faculty).

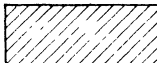
The problem was how to design a suitable calendar for student attendance. There was universal agreement on the fact that they should come for a normal freshman year of fall, winter and spring (see top of Table 5). And then the idea was that with twelve terms required for graduation, they would have nine more terms to choose out of twelve. There are 220 possible calendars, and that seems like an enormous amount of choice. Surprisingly, that answer turned out to be wrong.

In the figure 220 we have counted all possible combinations of how students could elect a schedule without taking into account whether they are reasonable. For example, we have counted the plan where after the freshman year the student takes off three terms and then goes nine terms in a row. Presumably no student will elect that plan. Therefore, some computer work was done (you can do it by hand, it just takes a long time), to try to put some conditions of reasonableness on the solutions, and to try to find out how many plans look at all attractive. The

TABLE 5 — DARTMOUTH PLAN

	Su	F	W	Sp
Frosh		×	×	×
Soph				
Junior				
Senior				

$(12_9) = 220$

	Su	F	W	Sp
Frosh		×	×	×
Soph			×	×
Junior	×	×		
Senior	×	×	×	×

rather surprising result was that starting with 220 possibilities, the number of reasonable plans was not high enough to make this plan acceptable. Of the roughly one dozen attractive plans the majority has the student on campus every winter term, which would not allow us to increase the enrollment.

The next proposed solution, and still my favorite, is not the one we have adopted. An excellent faculty-student committee came up with a single calendar plan and tried to persuade us that every student ought to go on this plan. (See the bottom of Table 5.) It's a marvelous scheme. It meets the boundary condition of a standard freshman year and the boundary condition of all students graduating in the spring of the senior year; it gives two six-month vacations, which is a very attractive option — students can get interesting jobs in that time period. It reduces the total number of terms from 12 to 11. And it has the marvelous feature that in fall, winter and spring only three of the four classes are present; therefore you can increase the total student body by one-third and still use the same space.

It was an extremely tempting plan, but the faculty did not buy it—for quite good reasons. The thought of going to something strange and have it totally required bothered the faculty, not to mention such shortcomings as the fact that there would be no sophomore football players, and no junior hockey and baseball players on campus. That's where reality impinges on combinatorics! (To be fair, the committee proposed a solution to this problem, but it was not a popular one.)

Nevertheless, the much greater choice under an 11-term plan led the faculty to a reduction of the graduation requirement to 33 courses. Since 32 courses is a typical requirement at many comparable schools, this was entirely reasonable. But the faculty opted for a maximum amount of freedom of choice for students. While the very attractive plan in Table 5 is a legal option, it is only one of many. Within certain limits each student would be allowed to design his own pattern for going through college. Every student will be required to come at least one summer, and we will reserve the right that if everyone wants to come in the same term, some students must take a second choice. The price of greater choice is that instead of increasing the student body by 33% we expect to increase it by about 25%.

I have shown you two numerical models and one combinatorial model. I would now like to show you a dynamic model, dealing with the use of endowment income. I have never seen this result in print, although once you have formulated the problem, the solution is quite trivial. The fact that such a basic result is not in general use may have something to do with the fact that the presence of a mathematician in the president's office is a rare event in American higher education.

Table 6 shows the three basic variables X , E and U , and six parameters. It is the interrelation of these quantities that I want to explore. Let us consider the six parameters. We assume an equilibrium situation in which these parameters remain constant. R_1 is the rate at which expenses grow annually, the fundamental quantity in budget control. R_2 is the rate at which you want your endowment to grow. R_3 is a crucial number, the total return you get on your endowment. In modern investment policy we don't care what comes in as dividends and what comes

TABLE 6 — ENDOWMENT USE

X	=	Expenses
E	=	Endowment
U	=	Endowment utilized
R_1	=	Rate of growth of expenses
R_2	=	Rate of growth of endowment
R_3	=	Rate of total return
R_4	=	% of endowment used
R_5	=	New endowment as % of old
R_6	=	% of expenses covered by endow.

in as capital growth; all you care about is the long run total return. With good investment management you should try to achieve something like 9% per year. R_4 is the percentage of the endowment that you use this year. R_5 is new endowment as a percentage of old endowment. Each year you get some new gifts. Although the amounts vary from year to year, the Dartmouth experience is that they represent a fairly steady percentage of your actual endowment, something like 2%. And finally R_6 is the percentage of the expenses covered by your endowment.

Table 7 shows a simple model containing the interrelations of the variables.

Equations (1)–(4) are, in effect, the definitions of R_1 , R_2 , R_4 and R_6 , respectively. Equation (5) is a balance sheet for endowment funds. The endowment next year will be the endowment this year plus the total return you get on it, to which you add new endowment that comes in in the form of gifts, and from which you subtract that which you have utilized. It is an absolutely straight-forward equation. What can one deduce from this simple model?

TABLE 7

- (1) $X_{n+1} = X_n \cdot (1 + R_1)$
 - (2) $E_{n+1} = E_n \cdot (1 + R_2)$
 - (3) $U_n = E_n \cdot R_4$
 - (4) $U_n = X_n \cdot R_6$
 - (5) $E_{n+1} = E_n \cdot (1 + R_3) + E_n \cdot (R_5) - U_n$
-
- (6) $X_n = X_0 \cdot (1 + R_1)^n$
 - (7) $E_n = E_0 \cdot (1 + R_2)^n$
 - (8) $E_0 R_4 \cdot (1 + R_2)^n = X_0 R_6 \cdot (1 + R_1)^n$
 - (9) $E_n \cdot (1 + R_2) = E_n (1 + R_3) + E_n \cdot R_5 - E_n \cdot R_4$

First of all, equations (1) and (2) lead to equations (6) and (7). We substitute these values into equations (3) and (4), setting the two values of U_n equal to each other, obtaining equation (8). If this is to hold for large n , we must have $R_1 = R_2$. That says that in the long run, if you want to get anything like an equilibrium situation, your endowment must continue to grow at the same rate that your expenses do. Obviously, for a few years you can violate this rule, but there is no escape in the long run.

A more interesting result is obtained from equation (5). We may use equations (2) and (3) to express all quantities in terms of E_n , as shown in (9). From this we deduce the very simple equation that $R_2 + R_4$ must, for any equilibrium solution equal $R_3 + R_5$. Here are the two major findings:

$$(10) \quad R_1 = R_2$$

$$(11) \quad R_1 + R_4 = R_3 + R_5.$$

In the second equation we have replaced R_2 by R_1 . Incidentally, the results are independent of R_6 , the percent of expenses covered by endowment. (This differs greatly from institution to institution; at Dartmouth it is about 1/3, at Harvard it seems to be 110%.)

These equations were derived from very simple-minded mathematics, but are most useful for long-range planning. The right-hand side of (11) contains your assets. R_3 is the rate of total return, say 9%, and R_5 is the percentage of new gifts, about 2%, adding up to about 11%. So you have to take a guess, if you are a trustee, as to what the sum of these two numbers will be and that tells you what you can

do. The left-hand side of (11) says that you must make an allocation of this available growth between R_2 the growth of endowment and R_4 the amount used currently. So here you have a conflict. You have a fixed total available to you and the more you use currently the less the endowment is going to grow, and the more it grows the less you can use currently. Put that way it is obvious, but, that it is a simple additive relationship, frankly surprises me.

The most fundamental decision for the institution is how to apply equation (11). At the present time Dartmouth estimates $R_3 + R_5 = 10.5\%$, and we are currently using $R_4 = 4.5\%$. This allows $R_1 = R_2 = 6\%$ for growth of expenses (and for the growth of the endowment). This is the kind of analysis that all boards of trustees ought to engage in, but to my knowledge they do not.

I would like to make a few remarks about the use of computers. Some of you may be surprised that I stayed away from computers this long in my talk. They are certainly not irrelevant to the job of the president. I personally, as many of you know, feel that the computer, particularly a good time-sharing system with its ease of access and very great ease of forming models, is an absolutely indispensable tool for education and I am beginning to see that it is also indispensable for college management.

Computer usage starts with absolutely trivial applications. I'll tell you a story of a nice opportunity to show off before the Trustees. It will be obvious to *you* that what I did was trivial, but it is *not* obvious to non-mathematicians. We were trying to figure out how fast tuition should increase and one of the Trustees asked how fast it has been increasing in the last decade. It was a reasonable question but none of us knew the answer. The Treasurer fortunately knew the present tuition and what tuition was ten years ago. That left the simple problem of finding out the rate of growth over ten years. You just take the ratio and extract a 10th root. I'm not going to say I did that in my head! But I happen to have a computer terminal in my office, and needed only a two-line program to solve the problem. So we interrupted a Trustee meeting for approximately 45 seconds so I could tell the Trustees that tuition at Dartmouth College had increased by exactly 6.3% per year, compounded, in ten years.

A second example is too complicated for detailed description in this talk. It is the question of how you design a loan plan for students, similar to that pioneered by Yale. Most people have a feeling that the Yale Plan is not quite right, but we have yet to demonstrate that a more attractive one exists. You run into some very complicated mathematical problems in designing a loan plan and computer simulation is almost indispensable. For example, everyone talks about how many years it will take until the plan "turns around", i.e., the repayments start exceeding the amount of money being borrowed. But it is far from obvious that the loan program ever turns around. On highly reasonable assumptions the amount borrowed increases faster than repayments. Or the debt outstanding is so high that the interest on it

makes catching up impossible. It is interesting to calculate what the maximum debt outstanding is; it is quite possible that if all universities in the country go on a reasonable loan plan at reasonable rates, then before the end of the century all the money in the world will be out in the form of loans to college students!

My final example is the use of computers for the dual purpose of a good management information system and for long-range planning models. I've had a great deal of frustration in that information is not available to decision-makers. As a faculty member I firmly believed that the administration was hiding facts from the faculty. After I got into office I found out that the facts were not available to the President either. The chances are that if you ask an office for factual information, they either can't find the facts or you are presented with a computer printout a foot high. I have a general rule that anybody who produces a computer printout that high must be doing the wrong thing. So the goal is to make information easily accessible and computer-readable.

For example, before I became President I served years on the committee on tenure and promotion. Each year we needed a list of associate professors who had been at least four years in rank, and wanted to know how long they had been in rank, what departments they were in, and how old they were. It seems like a modest enough request, but some poor secretary had to plow through a computer printout a foot high and spot associate professors, try to do a mental calculation as to whether they had been in rank four years, do another calculation on how old they were, and so on. Typically ten percent of the information was wrong. In one document the age was stated variously as 30, as 32 and as 35. I claim that a 5-year error on somebody who turned out to be 30 years old is more than three standard deviations! So our first goal is relatively modest, just to be able to lay one's hands on such information. If I had a computer terminal here I could show you that a good part of this is now working and is very helpful.

Dartmouth has a secret weapon in computer developments. Starting with the development of our original time-sharing system, since we couldn't afford to employ professionals, we have employed our own undergraduate students. We learned that they are often better than the professionals we could have hired. This has been our secret in developing one of the least expensive computing systems in the United States and, I think, one of the best. So when we went to a management information system, several faculty members worried about the design and five students did all the initial implementation. This is why in record time we have the beginnings of a management information system.

The second part of the project is going to be much harder and I don't really know what we are going to do. I feel a strong need for a model for long-range planning. I have looked at models at several other institutions and I'm sorry to say that I don't see their usefulness. They may be useful for justifying a budget to a legislature, but they have very little to do with real expenses.

The problem is that most models assume that expenses are linear and I haven't

yet found anything at Dartmouth that is linear. Let me give you the simplest example. In these models an important item is the cost of the library per student. I claim that the number of students has absolutely nothing to do with what the library costs! Dartmouth College happens to have a library with one million volumes in it, of which we are very proud. But I claim that if we had half as many students or twice as many students we would still have a collection of one million volumes. Perhaps the number of faculty members has somewhat more to do with costs, but only marginally. Basically, we have a million volumes because we're trying to maintain a first-rate library. Therefore any changes are really second order perturbations on that number.

So my big dilemma is how one builds a mathematical model which shows true interrelationships amongst parts of the institution, and not just in terms of money. What is the effect of new students on number of faculty members, and in turn what is the effect on space needs and how many more janitors we are going to need and how much more lawn has to be mowed, etc.? These questions are not even vaguely understood in academic life and yet someone has to come up with a suitable planning tool. Money is hard to get and to do well academically we must tighten up on our planning processes. But without rational tools I think we are going to make major goofs. Incidentally, if any of you have any good ideas on how to build such a model, I'd love to hear from you.

What conclusions can I draw from these examples? Let me start with a trite statement: planning the future of a modern university is a very complex enterprise, and the role of the president seems to be quite different from what it was 50 years ago. While I said at the beginning that for a great many activities being a mathematician is not relevant, there are a surprising number of mathematical-flavored problems that come up in the decision-making process. As I think you see from the examples, none of them require highly advanced mathematics—I didn't use homotopy theory once in any of my examples—for that matter I didn't even use calculus. Nevertheless, they are sufficiently complicated mathematically that the intelligent layman is not able to deal with them. Therefore, a college president either needs to be a mathematician or he needs someone first-rate on his planning staff who has considerable mathematical sophistication. And even if he has a mathematical planner, he has the handicap that he gets his information at second-hand.

When I've said all that, I have the feeling that I did not tell the full story of the relevance of mathematics for the college presidency. The college president spends an enormous amount of his time, when he is not doing junk, solving problems. I think if there is a single important thing that all college presidents should know about mathematics, it is how to think like a mathematician. A mathematician, or someone well trained in mathematics, has a special knack at analyzing problems and in knowing how to attack their solutions. Therefore, as a college president I am most grateful for being a mathematician.

SOME MATHEMATICAL VERSES

There once was a Norbert named
Wiener
Whose mind just couldn't be keener
But he'd chant and recite
Verses so trite
That we wished he'd sing less and
obscener.

* * *

A Valentine

My valentine I send to thee
With love and osculation
Dear love let's have no more harmonic
separation
Love's waves abound today
And I await thy reciprocation.

* * *

In the Greek mathematical Forum
Young Euclid was present to bore'em
He spent most of his time
Drawing circles sublime
And in crossing the *pons asinorum*.

* * *

A mathematician named Klein
Thought the Moebius band was divine
Said he, "If you glue
The edges of two
You'll get a weird bottle like mine."

* * *

Automation is vexation
Quaternions are bad
Analysis situs is detritus
I wonder, have I been had?

* * *

A quadratic function ambitious
Said it's not only wrong but it's vicious

It's surely no sin
To have both max and min
To limit me so is malicious.

* * *

Here's a toast to L. J. Mordell
Young in spirit, most active as well
He'll never grow weary
Of his love, number theory
The results that he gets are just swell.

* * *

It used to be fun
To add one and one
But now I'm unsure
What sum to secure
I'm told it may even be none.

* * *

There once was a function of x
With deplorable notions of sex
In a half-Baire condition
It attempted coition
With a function weakly convex.

* * *

Here's a toast to Leo H. Moser
A truly delightful exposé
Of things mathematical
With a flair for dramatical
With him you cannot be a poser.

* * *

A question both real and complex
Is whether a sphere is convex
That problem proposer
The famed Leo Moser
Believes it depends on the sex.
(S. Golomb)

* The verses here were found in papers left by the late Leo Moser. They were communicated to the Monthly by his brother, William Moser, of McGill University.

REMARKS ON WOMEN IN MATHEMATICS

ASSOCIATION FOR WOMEN IN MATHEMATICS, Philadelphia Chapter, Temple University

As the Philadelphia Chapter of the Association for Women in Mathematics, we feel it is not only appropriate but necessary to respond to the article by University of Pennsylvania professor Murray Gerstenhaber in the June-July issue of the MONTHLY, 1972.

The article presents stereotyped, derogatory and negative views of women in the mathematical, academic and professional world. Professor Gerstenhaber seems to regard childrearing as an insurmountable obstacle to women's intellectual and professional achievement. He completely ignores the fact that many women continuously pursue challenging careers and make significant contributions to their chosen professional fields. He would have women channeled into careers in recently mathematicized fields where their "principal expertise (would consist) in using a computer cleverly." The reasoning is that they can continue such careers from their homes, to which they are presumably tied by maternal duties. But women are by no means so restricted.

Many women continue working while their children are small.¹ Many who do interrupt their careers to raise children want an alternative. But why relatively undemanding work to be done at home? To propose this as the only course open to them is to ignore large changes going on in society. Childcare centers are coming more and more to be perceived as potentially enriching to the child as well as liberating to the parents.² More and more couples, especially in academia, find that by sharing family responsibilities both are able to pursue their careers.³ In many quarters the traditional, confining concept of woman's role is being challenged.

In light of this, we feel Professor Gerstenhaber has made the wrong prognosis, not only about the careers women will choose, but also about the appropriate response for mathematics departments to the expected influx of women. His solution is to "omit proofs," that is, to lower standards.

Mathematics departments have always functioned to teach skills to students in other disciplines, e.g., engineering. Why is it now, with women anticipated in large numbers, that we must suddenly begin to omit proofs? The implication, intended by the author or not, is that women can't do mathematics as well as men. Another implication scarcely avoidable when computer-oriented careers are advocated for women is that it is a good choice for them because of its routine nature, the work being often that of a technologically trained secretary, and again women are put in the same limited slot of accessory that they've always been in.

The math students of the future may be predominantly women, but we cannot foresee the necessity for a "freshman course . . . taught like arithmetic in grade school." There are contributions to be made in advanced theoretical research, and some, we

are sure, will be made by women who have been taught something other than the techniques of computer programming.

Mathematics departments may be expected to offer courses to those of both sexes who desire only to be trained for computer applications, but we feel that no self-respecting mathematics department will ever feel the need to lower its standards in order to educate its women students.

NOTES

1. For example, of a group of 1957–58 Radcliffe Ph.D.s, 91 % were employed in 1964 and 79 % had never had any interruption in the full-time pursuit of their professions. The time period of the study corresponded to the child rearing years of most women. See Patricia A. Graham, "Women in Academe," *Science*, Vol. 169, No. 3995, 1284–90.

2. It is now widely accepted that children do not suffer if their mothers work. Rather it is the quality of attention the child receives which influences his or her emotional development. See Alice Rossi, "Women in Science, Why so few?" *Science*, Vol. 148, No. 3674, 1196–1202.

3. "Marriages in Academia Reflect the Changing Status of Women," *New York Times*, Nov. 13, 1972.

MATHEMATICAL NOTES

EDITED BY ROBERT GILMER

Material for this Department should be sent to David Roselle, Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803.

A NOTE ON CATALAN PARENTHESES

JOHN RIORDAN, The Rockefeller University

E. Catalan [1], in 1838, proposed and solved the problem of finding the number of ways of evaluating the product of n factors (in fixed order) by successive multiplications operating always on two adjacent factors; the number in question is c_{n-1} , where

$$c_n = (n + 1)^{-1} \binom{2n}{n}.$$

Accounts of this work appear in [2] and [5]. For $n = 4$, the five ways are:

$$((ab)(cd)), (((ab)c)d), ((a(bc))d), (a((bc)d)), (a(b(cd))).$$

The first of these differs from the others in having two multiplications of the given factors; the other four, following [3] and [4] may be called "nests", and in the same

spirit the Catalan parentheses are “clutches of nests”. What is the number of Catalan parentheses of n factors with k nests?

Write $c_n(x)$ for the enumerator of clutches by number of nests (the coefficient of x^k in $c_n(x)$, $c_{n,k}$, is the number of parentheses of n factors with k nests). Then, following Catalan,

$$(1) \quad c_n(x) = c_1(x)c_{n-1}(x) + \cdots + c_j(x)c_{n-j}(x) + \cdots + c_{n-1}(x)c_1(x)$$

with initial values $c_1(x) = 1$, $c_2(x) = x$. If $c(x, y) = \sum_{n=1} y^n c_n(x)$, it follows from (1) and the initial conditions that

$$(2) \quad \begin{aligned} c(x, y) &= y + xy^2 + yc_1(x)(c(x, y) - yc_1(x)) + y^2c_2(x)c(x, y) + \cdots \\ &+ y^jc_j(x)c(x, y) + \cdots = y + (x-1)y^2 + c^2(x, y). \end{aligned}$$

The solution of (2) satisfying the initial conditions is

$$(3) \quad c(x, y) = \frac{1}{2}[1 - (1 - 4y - 4(x-1)y^2)^{\frac{1}{2}}]$$

or

$$(3a) \quad c(x, y) = (y + (x-1)y^2)c(y + (x-1)y^2),$$

with $2yc(y) = 1 - \sqrt{1-4y}$; $c(y)$ is the generating function of the Catalan numbers c_n , defined above.

Expansion of (3a) leads at once to

$$(4) \quad c_n(x) = \sum_{k=0} \binom{n-k}{k} (x-1)^k c_{n-1-k}, \quad n = 1, 2, \dots$$

As a verification, note that by (4) $c_n(1) = c_{n-1}$.

Further results seem to flow more easily from (2). First, denoting partial derivatives of $c(x, y)$ by suffixes, it follows from (2) that

$$(5) \quad \begin{aligned} [1 - 2c(x, y)]c_x(x, y) &= y^2, \\ [1 - 2c(x, y)]c_y(x, y) &= 1 + 2(x-1)y. \end{aligned}$$

Hence

$$(6) \quad [1 + 2(x-1)y]c_x(x, y) = y^2c_y(x, y)$$

and, with primes denoting derivatives,

$$(7) \quad c'_n(x) + 2(x-1)c'_{n-1}(x) = (n-1)c_{n-1}(x)$$

which implies

$$(8) \quad kc_{n,k} = 2kc_{n-1,k} + (n+1-2k)c_{n-1,k-1}.$$

The first few values of $c_n(x)$ are: $c_1(x) = 1$, $c_2(x) = x$, $c_3(x) = 2x$, $c_4(x) = 4x + x^2$.

It is apparent that $c_{n0} = \delta_{n1}$ (Kronecker delta), and that the degree of $c_n(x)$ is the integral part of $n/2$.

The instance $k = 1$ of (8) and $c_{n0} = \delta_{n1}$, show that $c_{21} = 1$ and $c_{n1} = 2c_{n-1,1}$, $n = 2, 3, \dots$; hence $c_{n1} = 2^{n-2}$, $n = 2, 3, \dots$. Similarly it is found that

$$c_{n2} = \binom{n-2}{2} 2^{n-4}, \quad n = 4, 5, 6, \dots$$

which leads to the guess

$$c_{nk} = \binom{n-2}{2k-2} 2^{n-2k} c_{k-1}$$

which satisfies (8) and the boundary conditions.

Thus finally

$$(9) \quad c_n(x) = \sum_{k=1}^m \binom{n-2}{2k-2} 2^{n-2k} c_{k-1} x^k, \quad m = \left\lfloor \frac{n}{2} \right\rfloor, \quad n = 2, 3, \dots$$

An interesting aspect of (9) is the identity

$$(10) \quad c_{n+2}(1) = c_{n+1} = \sum_{k=0}^m \binom{n}{2k} 2^{n-2k} c_k$$

given by Jacques Touchard [6], in 1928, and now acquiring a combinatorial meaning.

References

1. E. Catalan, Note sur une équation aux différences finies, J. Math. Pures Appl., (1) 3 (1838) 508–516.
2. Louis Comtet, Analyse Combinatoire, Paris 1970; Tome Premier, p. 64 ff.
3. G. Eldredge, Nesting habits of the ladder parenthesis, Problem E 1903, this MONTHLY, 73 (1966) 666; M. Goldberg, incomplete solution, *ibid.* 77 (1970) 525–526; E. F. Schmeichel, comment, *Ibid.* 78 (1971) 298.
4. Richard K. Guy and J. L. Selfridge, The nesting and roosting habits of the ladder parenthesis, Research Paper No. 127 (June 1971), The University of Calgary; see also this issue page 868.
5. E. Netto, Lehrbuch der Combinatorik, Chelsea, New York, 1958, p. 192 ff.
6. J. Touchard, Sur certaines équations fonctionnelles, Proc. Internat. Congr. Math., University of Toronto Press, Toronto, 1928, 465–472.

DETERMINATION OF THE RIEMANN FUNCTION

E. J. SCOTT, University of Illinois

1. Introduction. The Riemann method is often used to solve a Cauchy problem relative to the hyperbolic equation

$$(1.1) \quad \frac{\partial^2 u}{\partial y \partial x} + \alpha(x, y) \frac{\partial u}{\partial x} + \beta(x, y) \frac{\partial u}{\partial y} + \gamma(x, y) u = \delta(x, y),$$

where α, β, γ are assumed twice continuously differentiable functions in some region R and δ is continuous there. To do this one must first find the Riemann function $v(x, y)$ which satisfies the adjoint equation of (1.1), namely,

$$(1.2) \quad \frac{\partial^2 v}{\partial y \partial x} - \frac{\partial}{\partial x}(\alpha v) - \frac{\partial}{\partial y}(\beta v) + \gamma v = 0,$$

$$x \geq 0, y \geq 0,$$

the boundary conditions

$$(1.3) \quad \frac{dv(x, 0)}{dx} - \beta(x, 0)v(x, 0) = 0,$$

$$(1.4) \quad \frac{dv(0, y)}{dy} - \alpha(0, y)v(0, y) = 0,$$

and the condition

$$(1.5) \quad v(0, 0) = 1.$$

This auxiliary problem, which is known as a Goursat problem, has a unique solution under the assumptions made above [2].

2. Use of the Laplace transformation. We shall show how, under appropriate conditions to be given subsequently, it is possible to solve the Goursat problem for the Riemann function given by equations (1.2) to (1.5) by means of the Laplace transformation [1]. To that end, let us take the Laplace transform of equation (1.2) with respect to x where

$$L\{v(x, y)\} = \int_0^\infty e^{-sx} v(x, y) dx = \bar{v}(s, y).$$

Then, since

$$L\left\{\frac{\partial^2 v}{\partial y \partial x}\right\} = s \frac{d\bar{v}(s, y)}{dy} - \frac{dv(0, y)}{dy},$$

$$L\left\{\frac{\partial}{\partial x}(\alpha v)\right\} = s \int_0^\infty e^{-sx} (\alpha v) dx - \alpha(0, y)v(0, y),$$

(assuming that $\operatorname{Re}(s) > 0$ and α, v are such that $\lim_{x \rightarrow \infty} (\alpha v) e^{-sx} = 0$),

$$L\left\{\frac{\partial}{\partial y}(\beta v)\right\} = \frac{\partial}{\partial y} \int_0^\infty e^{-sx} (\beta v) dx,$$

(assuming the validity of interchanging the operations of integration and differentiation), and

$$L\{\gamma v\} = \int_0^\infty e^{-sx} (\gamma v) dx,$$

equation (1.2) is transformed into

$$s \frac{d\bar{v}(s, y)}{dy} - s \int_0^\infty e^{-sx}(\alpha v)dx - \frac{\partial}{\partial y} \int_0^\infty e^{-sx}(\beta v)dx \\ + \int_0^\infty e^{-sx}(\gamma v)dx - \left[\frac{dv(0, y)}{dy} - \alpha(0, y)v(0, y) \right] = 0,$$

or, in view of condition (1.4),

$$(1.6) \quad s \frac{d\bar{v}(s, y)}{dy} - s \int_0^\infty e^{-sx}(\alpha v)dx \\ - \frac{\partial}{\partial y} \int_0^\infty e^{-sx}(\beta v)dx + \int_0^\infty e^{-sx}(\gamma v)dx = 0.$$

Because of the presence of the products αv , βv , and γv appearing in the integrands of the last three terms, progress in determining the function $\bar{v}(s, y)$ is somewhat limited unless we make some further simplifying, but not unduly restrictive assumptions on the functions α, β, γ . To obtain a tractable equation involving $\bar{v}(s, y)$, we shall assume that α, β, γ are functions of y alone. If this be so, then equation (1.6) becomes the first order linear differential equation

$$(1.7) \quad [s - \beta(y)] \frac{d\bar{v}(s, y)}{dy} + [\gamma(y) - \beta'(y) - s\alpha(y)]\bar{v}(s, y) = 0$$

whose solution is readily shown to be

$$(1.8) \quad \bar{v}(s, y) = c \exp \left[\int_0^y \alpha(\tau) d\tau \right] \cdot \exp \left[- \int_0^y \frac{\gamma(\tau) - \beta'(\tau) - \alpha(\tau)\beta(\tau)}{s - \beta(\tau)} d\tau \right],$$

where c is generally a function of s . To determine c we turn to condition (1.3) and note that if β is a function of y alone it becomes the differential equation

$$(1.9) \quad \frac{dv(x, 0)}{dx} - \beta(0)v(x, 0) = 0$$

the Laplace transform of which is

$$s\bar{v}(s, 0) - v(0, 0) - \beta(0)\bar{v}(s, 0) = 0,$$

or, in view of (1.5),

$$(1.10) \quad \bar{v}(s, 0) = \frac{1}{s - \beta(0)}.$$

We conclude from equations (1.8) and (1.10) that

$$(1.11) \quad c = \frac{1}{s - \beta(0)}.$$

Hence,

$$\bar{v}(s, y) = \exp \left[\int_0^y \alpha(\tau) d\tau \right] \frac{\exp \left[- \int_0^y \frac{\gamma(\tau) - \beta'(\tau) - \alpha(\tau)\beta(\tau)}{s - \beta(\tau)} d\tau \right]}{s - \beta(0)},$$

whence,

$$(1.12) \quad v(x, y) = \exp \left[\int_0^y \alpha(\tau) d\tau \right] L^{-1} \left\{ \frac{\exp \left[- \int_0^y \frac{\gamma(\tau) - \beta'(\tau) - \alpha(\tau)\beta(\tau)}{s - \beta(\tau)} d\tau \right]}{s - \beta(0)} \right\}.$$

From this formula we can obtain solutions to a number of specific problems. For example: If $\gamma(y) = \alpha(y)\beta(y)$, then

$$v(x, y) = \exp \left[\int_0^y \alpha(\tau) d\tau \right] L^{-1} \left\{ \frac{1}{s - \beta(y)} \right\} = e^{x\beta(y)} \exp \left[\int_0^y \alpha(\tau) d\tau \right].$$

If $\gamma(y) \neq \alpha(y)\beta(y)$ and $\alpha(y), \beta(y), \gamma(y)$ are the constants α, β, γ , respectively, then

$$v(x, y) = e^{xy} L^{-1} \left\{ \frac{\exp \left[- \frac{\gamma - \alpha\beta}{s - \beta} \right] y}{s - \beta} \right\} = e^{xy + \beta x} J_0(2\sqrt{(\gamma - \alpha\beta)xy}).$$

In particular, if

- (a) $\alpha = \beta = \gamma = 0$, then $v(x, y) = 1$,
- (b) $\alpha = 0, \beta = 0, \gamma \neq 0$, then $v(x, y) = J_0(2\sqrt{\gamma xy})$,
- (c) $\alpha = 0, \beta \neq 0, \gamma \neq 0$, then $v(x, y) = e^{\beta x} J_0(2\sqrt{\gamma xy})$,
- (d) $\alpha \neq 0, \beta = 0, \gamma \neq 0$, then $v(x, y) = e^{xy} J_0(2\sqrt{\gamma xy})$.

As another example, let $\alpha(y) = -1$, $\beta(y) = y$ and $\gamma(y) = 2 - y$. Then, from (1.12)

$$\begin{aligned} v(x, y) &= e^{-y} L^{-1} \left\{ \frac{\exp \left[- \int_0^y d\tau / (s - \tau) \right]}{s} \right\} \\ &= e^{-y} L^{-1} \{ (s - y) / s^2 \} = e^{-y} (1 - xy). \end{aligned}$$

Remark. A similar analysis can be carried out if $\alpha(x, y), \beta(x, y), \gamma(x, y)$ be assumed functions of x alone. The Laplace transform is then taken with respect to y .

References

1. G. Doetsch, *Theorie und Anwendung der Laplace-Transformation*, Springer-Verlag, Berlin, 1937.
2. A. N. Tychonov and A. A. Samarski, *Partial Differential Equations of Mathematical Physics*, Vol. 1, Holden-Day, San Francisco, 1964, page 114.

INEQUALITIES FOR THE AREA OF TWO TRIANGLES

L. CARLITZ, Duke University

Let a, b, c denote the sides of the triangle ABC and let a', b', c' denote the sides of the triangle $A'B'C'$. Let F, F' denote the respective areas. Pedoe [2] has proved that

$$(1) \quad a^2(-a'^2 + b'^2 + c'^2) + b^2(a'^2 - b'^2 + c'^2) + c^2(a'^2 + b'^2 - c'^2) \geq 16FF',$$

with equality if and only if the triangles $ABC, A'B'C'$ are similar. The writer [1] has recently given a simple algebraic proof of (1).

Since

$$F^2 + F'^2 \geq 2FF',$$

it is natural to ask whether

$$(2) \quad H \geq 8(F^2 + F'^2),$$

where H denotes the left hand side of (1). We shall prove the following result.

THEOREM. *If the differences*

$$(3) \quad a^2 - a'^2, \quad b^2 - b'^2, \quad c^2 - c'^2$$

are all positive or all negative and in addition the numbers

$$(4) \quad |a^2 - a'^2|^{\frac{1}{3}}, \quad |b^2 - b'^2|^{\frac{1}{3}}, \quad |c^2 - c'^2|^{\frac{1}{3}}$$

form the sides of a triangle \triangle (possibly degenerate) then

$$(5) \quad 8(F^2 + F'^2) - H = 8F^2(\triangle),$$

where $F(\triangle)$ denotes the area of \triangle . Otherwise

$$(6) \quad H \geq 8(F^2 + F'^2).$$

Proof. Since

$$16F^2 = 2b^2c^2 + 2c^2a^2 + 2a^2b^2 - a^4 - b^4 - c^4,$$

$$16F'^2 = 2b'^2c'^2 + 2c'^2a'^2 + 2a'^2b'^2 - a'^4 - b'^4 - c'^4,$$

it is easily verified that

$$(7) \quad 16(F^2 + F'^2) - 2H = 2 \sum (a^2 - a'^2)(b^2 - b'^2) - \sum (a^2 - a'^2)^2.$$

Now assume that the differences (3) are all positive or all negative and put

$$x = |a^2 - a'^2|^{\frac{1}{3}}, \quad y = |b^2 - b'^2|^{\frac{1}{3}}, \quad z = |c^2 - c'^2|^{\frac{1}{3}},$$

so that (7) becomes

$$(8) \quad 16(F^2 + F'^2) - 2H = 2 \sum x^2 y^2 - \sum x^4.$$

We recall that three positive numbers x, y, z form the sides of a triangle if and only if

$$(9) \quad 2 \sum x^2 y^2 - \sum x^4 \geq 0.$$

Moreover, the left hand side of (9) is equal to sixteen times the square of the area of this triangle. This evidently proves (5).

If however $2 \sum x^2 y^2 - \sum x^4 < 0$, it follows from (8) that

$$(10) \quad H > 8(F^2 + F'^2).$$

If $z = 0$, say, (8) reduces to

$$16(F^2 + F'^2) - 2H = 2x^2 y^2 - x^4 - y^4 < 0,$$

(unless $x = y = 0$, so that (10) holds in this case also).

Finally assume that not all the differences are of the same sign. We may suppose that

$$a^2 - a'^2 = -u, \quad b^2 - b'^2 = v, \quad c^2 - c'^2 = w,$$

where $u \geq 0, v \geq 0, w \geq 0$. Then (7) becomes

$$\begin{aligned} 16(F^2 + F'^2) - 2H &= -2uv - 2uw + 2vw - u^2 - v^2 - w^2 \\ &= -(u + v - w)^2 - 4uw \leq 0, \end{aligned}$$

with equality only when $u = 0, v = w$.

This completes the proof of the theorem.

Supported in part by NSF grant GP-17031.

References

1. L. Carlitz, An inequality involving the area of two triangles, this MONTHLY, 78 (1971) 772.
2. D. Pedoe, Thinking geometrically, this MONTHLY, 77 (1970) 711-721.

A BANACH SPACE CHARACTERIZATION OF THE SPACE OF AFFINE CONTINUOUS FUNCTIONS ON A COMPACT CONVEX SET

P. D. TAYLOR, Queen's University at Kingston, Ontario, Canada

A subset K of a linear space E is called **convex** if K contains, with any two of its points, the line segment joining them. A real-valued function f on K is called **affine**

if $\text{graph}(f)$ is convex. For example, any linear functional on E is, when restricted to K , affine. Algebraically we have

$$K \text{ convex} \Leftrightarrow x, y \in K \Rightarrow tx + (1-t)y \in K \quad \forall t, 0 \leq t \leq 1,$$

$$f \text{ affine} \Leftrightarrow f(tx + (1-t)y) = tf(x) + (1-t)f(y) \quad \forall t, 0 \leq t \leq 1.$$

Now let's put topology in the picture. Let E be a linear *topological* space, K a *compact*, convex subset and let $A(K)$ denote the set of *continuous* affine functions on K . With the sup norm

$$\|f\| = \sup\{|f(x)| : x \in K\},$$

$A(K)$ is a real Banach space, that is, a real, complete, normed, linear space.

Now ask the following question: *Suppose we are given a real Banach space. How can we tell whether or not it is the space of affine continuous functions on some compact convex set?*

Let's be precise. The correct notion of equivalence for Banach spaces is *isometry*. Two Banach spaces are **isometric** if there is a bijective linear map Φ between them which preserves the norms: $\|\Phi(a)\| = \|a\|$. Such a map is called an **isometry**. Now we can state our problem precisely.

Given a real Banach space, find a property of the unit ball which will allow us to tell whether or not the space is isometric to the space of affine continuous functions on some compact convex set?

Notice that we have asked for a characterizations in terms of the *unit ball*. The **unit ball** is the set of elements of norm less than or equal to one and it determines and is determined by the norm. It is often more illuminating to think in terms of the unit ball rather than the norm. For example, an isometry is just a bijective linear map which sends the unit ball of one space onto the unit ball of the other.

Let's look at an example. Suppose K is an interval $[a, b]$. A member of $A(K)$ is determined by its values at the end points of K . In fact the map $f \rightarrow (f(a), f(b))$ is a linear bijection between $A(K)$ and \mathbb{R}^2 . If we give \mathbb{R}^2 the sup norm, $\|(x, y)\| = \max\{|x|, |y|\}$, this map will be an isometry.

Actually whenever $A(K)$ is two-dimensional, K is an interval. In general, if K contains n affinely independent points, then clearly $\dim A(K) \geq n$. So if $\dim A(K) = 2$ then K must be a compact convex subset of the real line, hence a closed interval.

Now let's solve our problem for two-dimensional Banach spaces. We have shown that any two-dimensional $A(K)$ space is isometric to \mathbb{R}^2 with the sup norm. What other norms can \mathbb{R}^2 have, which are isometric to the sup norm? The unit ball of the sup norm is a square, so in terms of unit balls we are asking: what is the image of a square under a bijective linear transformation of \mathbb{R}^2 ? A parallelogram, of course! Let us define a **parallelogram** in a Banach space to be the convex hull of four points u, v, w and x , not all in a straight line, for which $u + v = w + x$.

Then a two-dimensional Banach space is isometric to some $A(K)$ if and only if the unit ball is a parallelogram. Indeed since any bijective linear image of a parallelogram is a parallelogram, we have shown that the condition is necessary. On the other hand, any Banach space whose unit ball is a parallelogram, is by the obvious map, isometric to R^2 with the sup norm, hence to $A([0, 1])$.

For another example suppose K is a triangle with vertices a , b and c . Then $A(K)$ is the set of functions with planar graphs, and the map $f \rightarrow (f(a), f(b), f(c))$ is a linear bijection between $A(K)$ and R^3 . Since $|f|$ must attain its maximum at a vertex of K , this map is an isometry if we give R^3 the sup norm. This time the unit ball of R^3 is a cube

But now in three dimensions it is no longer true that any two $A(K)$ spaces are isometric. For example, take K to be the square with vertices a , b , c and d (in cyclic order). Then $A(K)$ is again the set of functions with planar graphs and the map $f \rightarrow (f(a), f(b), f(c))$ is a linear bijection between $A(K)$ and R^3 . But the sup norm on R^3 will not make this an isometry, since $|f|$ need not attain its maximum at a , b or c . Since $f((a+c)/2) = f((b+d)/2)$ we have $f(d) = f(a) + f(c) - f(b)$ (since f affine), and hence the above map is an isometry if we give R^3 the norm

$$\|(x, y, z)\| = \max\{|x|, |y|, |z|, |x + z - y|\}.$$

With this norm, the unit ball of R^3 is no longer the whole cube, but only that part of the cube between the two planes $x + z - y = \pm 1$. We can visualize this unit ball by observing that the plane $x + z - y = 1$ cuts off the vertex $(1, -1, 1)$ from the cube by passing through the three adjacent vertices, $(-1, -1, 1)$, $(1, 1, 1)$ and $(1, -1, -1)$. Similarly the plane $x + z - y = -1$ cuts off the opposite vertex $(-1, 1, -1)$ by passing through its three adjacent vertices. This is certainly not a parallelepiped, so that this norm is not isometric to the sup norm.

Let us recall our problem: to characterize the unit balls of $A(K)$ spaces. The unit ball of the last example, although not a parallelepiped, is still "parallelogramic" in nature in the following sense: any plane passing through the vertices $(1, 1, 1)$ and $(-1, -1, -1)$ cuts the unit ball in a parallelogram. This property can be translated via the isometry to the $A(K)$ space. The vertex $(1, 1, 1)$ corresponds to the constant function 1 and thus if K is a square, any two-dimensional subspace of $A(K)$ containing the constant function 1 intersects the unit ball in a parallelogram. We can ask if this is true for any $A(K)$ space. The answer is yes, and furthermore, this property characterizes $A(K)$.

THEOREM. *Let A be a Banach space with unit ball B . Then A is isometric to $A(K)$ for some compact convex K if and only if there is a point e in B with the property that every two-dimensional subspace of A containing e intersects B in a parallelogram with e as a vertex.*

Proof. If two spaces are isometric and one of them has the condition, so does the other. So to prove the necessity it is enough to show that $A(K)$ has the condition. Suppose K is compact, convex. Let e be the constant function 1 on K . If L is a two-dimensional subspace of $A(K)$ containing e , choose $f \in L$ such that $\|f\| = 1$, $\sup_K(f) = 1$ and $f \neq e$. Let $m = \inf_K f$. Then $m < 1$ since $f \neq e$. Let

$$g = \frac{2}{(1-m)} \left(f - \frac{(m+1)}{2} e \right).$$

Then $g \in L$ and an easy calculation shows $\sup_K g = 1$ and $\inf_K g = -1$. Hence the parallelogram with vertices $\pm e$ and $\pm g$ has the property that every point on its boundary has norm 1. This parallelogram must then be the unit ball of L .

To prove the converse we need to know a bit about Banach spaces. If A is a Banach space we denote by A^* the linear space of norm-continuous linear functionals on A . With the norm

$$\|\gamma\| = \sup\{|\gamma(a)| : a \in A, \|a\| \leq 1\} \quad (\gamma \in A^*),$$

A^* is a Banach space. By the **weak * topology** on A^* , we mean the weakest topology in which the elements of A , considered as functionals on A^* , are continuous. With this topology A^* is a linear topological space. A discussion of this topology can be found in Royden [3]. One of the basic results is that the unit ball of A^* is compact in the weak * topology [3, Chap. 10, Theorem 15].

Now suppose A is a Banach space, and e is a point in the unit ball B with the property of the theorem. If $\dim A = 1$ then A is isometric to $A(K)$ for any K containing only one point, and we are finished. So suppose $\dim A \geq 2$. Then A contains a two-dimensional subspace L such that e is a *vertex* of the parallelogram $L \cap B$. It follows that $\|e\| = 1$. Let

$$K = \{\gamma \in A^* : \|\gamma\| = 1 \text{ and } \gamma(e) = 1\}.$$

First notice that K is a weak * closed subset of B^* , the unit ball of A^* . Indeed, any γ in the closure of K must have $\gamma(e) = 1$ and $\|\gamma\| \leq 1$. Since $\|e\| = 1$ we have $\|\gamma\| \geq 1$, and hence $\gamma \in K$. Since B^* is weak * compact, so is K . Notice next that K is convex. Indeed any γ on the line segment between two points of K must have $\gamma(e) = 1$ and $\|\gamma\| \leq 1$. Again since $\|e\| = 1$ we have $\|\gamma\| \geq 1$ and $\gamma \in K$. So that K is a compact convex subset of a linear topological space.

Since members of A , if considered as functionals on A^* are linear and weak * continuous, we have a natural linear map $\Phi: A \rightarrow A(K)$, by restricting members of A to K . For $a \in A$ let

$$\|a\|_K = \sup\{\gamma(a) : \gamma \in K\}.$$

We shall have shown that Φ preserves norm if we show that $\|a\| = \|a\|_K$. It will

be enough to show this whenever $\|a\| = 1$ and $a \neq \pm e$. In this case let L be the two-dimensional subspace of A spanned by a and e . Since the unit ball of L is a parallelogram with e as a vertex, and a in the boundary, we can find $\lambda \in L^*$ such that $\lambda(e) = 1$, $|\lambda(a)| = 1$ and $\|\lambda\| = 1$. Use the Hahn-Banach Theorem [3, Chap. 10, Thm. 4], to choose $\gamma \in A^*$ with norm 1 and extending λ . Then $\gamma \in K$ and $|\gamma(a)| = 1$ from which $\|a\|_K \geq 1$. Since $K \subset B^*$, $\|a\|_K \leq 1$ and $\|a\|_K = 1 = \|a\|$. Thus Φ preserves norm, and is also injective.

It remains to show that Φ is onto. It will be enough to show that $\Phi(A)$ is norm-dense in $A(K)$. Indeed A is complete and Φ preserves norm, so that $\Phi(A)$ is complete and hence norm-closed in $A(K)$.

To show that $\Phi(A)$ is norm-dense in $A(K)$, we need two results from linear topological spaces. First, if M is such a space and we give its dual space M^* the weak * topology, what is the space of continuous linear functionals on M^* ? Clearly any member of M , regarded as a function on M^* , is weak *continuous (by definition of the weak * topology!). The result is that there are no others: *the space of weak *continuous linear functionals on M^* is precisely M* . The proof is short, but tricky, and can be pieced together from Kelly, Namioka [1, 17.6]. Secondly we need to know that if K is a compact convex subset of a linear topological space E then the linear span of the constants and the members of E^* , restricted to K , is norm dense in $A(K)$. This is an application of the separating-hyperplane theorem, and can be found in Phelps [2, Prop. 4.5].

Now we have what we need. By the first result, A is the dual space of A^* with the weak * topology. Since K is a compact convex subset of A^* with this topology, the second result tells us that $\Phi(A)$ together with the constants, spans a norm-dense subspace of $A(K)$. Since $\Phi(A)$ already contains the constants, we are finished.

References

1. J. L. Kelly and I. Namioka, *Linear Topological Space*, Van Nostrand, Princeton, N. J., 1963.
2. R. R. Phelps, *Lectures on Choquet's Theorem*, Van Nostrand, Princeton, N. J., 1966.
3. H. L. Royden, *Real Analysis*, Macmillan, New York, 1966.

ON LATTICE POINTS INSIDE CONVEX BODIES

A. M. ODLYZKO, California Institute of Technology

If K is a convex body in n -dimensional Euclidean space, let $V(K)$, $A(K)$, and $N(K)$ denote, respectively, the n -dimensional volume of K , the $(n-1)$ -dimensional surface area of K , and the number of lattice points (points with integer coordinates) in the interior of K . For each positive real number α , let

$$l(\alpha, n) = \min\{N(K): V(K)/A(K) \geq \alpha\}.$$

Hadwiger [2] has recently proved that for all $n \geq 2$ we have

$$(1) \quad l(\tfrac{1}{2}, n) \geq 1,$$

while

$$l(\alpha, n) = 0 \text{ for all } \alpha < \tfrac{1}{2}.$$

Essentially the same result had been proved previously by Bender [1] for $n = 2$, and Wills [5], [6] has shown that (1) holds for $n = 2, 3$, and 4.

It follows from (1) and the superadditivity of $l(\alpha, n)$ as a function of α that

$$(2) \quad l(\alpha, n) \geq [2\alpha] \text{ for } n \geq 2,$$

where $[y]$ is the greatest integer $\leq y$ (see [3] and [7]). This lower bound was raised in the two-dimensional case by Hammer [4], who used Bender's result to show that if r is a non-negative integer, then

$$l(2^r, 2) \geq 2^{r+2} - 1.$$

Combined with the superadditivity of $l(\alpha, n)$ that result gives an improved estimate of $l(\alpha, 2)$ for general α . This note proves the much stronger and more general result

$$(3) \quad l(\alpha, n) \geq [2\alpha]^n \text{ for } n \geq 2.$$

We also show that

$$(4) \quad l(\alpha, n) \leq \frac{\pi^{n/2}}{\Gamma((n/2) + 1)} (n\alpha + \sqrt{n})^n \text{ for } n \geq 1,$$

where Γ is the gamma function.

To prove (3) consider any convex body K such that $V(K)/A(K) \geq \alpha$, and let $r = [2\alpha]$. We can assume that $r \geq 1$, since otherwise (3) is trivial. Define

$$K' = \frac{1}{r}K = \left\{ \frac{1}{r}\mathbf{x} : \mathbf{x} \in K \right\}.$$

Then

$$V(K') = \frac{1}{r^n}V(K) \text{ and } A(K') = \frac{1}{r^{n-1}}A(K).$$

Now let $\mathbf{a} = (a_1, \dots, a_n)$ be any n -tuple of integers such that $0 \leq a_i \leq r-1$ for $i = 1, \dots, n$ and consider

$$K' - \frac{1}{r}\mathbf{a} = \left\{ \mathbf{z} - \frac{1}{r}\mathbf{a} : \mathbf{z} \in K' \right\}.$$

Since $K' - (1/r)\mathbf{a}$ is a translate of K' , it has the same volume and surface area as K' . In particular

$$\frac{V(K' - (1/r)\mathbf{a})}{A(K' - (1/r)\mathbf{a})} = \frac{V(K')}{A(K')} = \frac{1}{r} \frac{V(K)}{A(K)} \geq \frac{\alpha}{[2\alpha]} \geq \frac{1}{2}.$$

Hadwiger's theorem (inequality (1)) now asserts that $K' - (1/r)\mathbf{a}$ contains at least one lattice point. This means that K' contains a point of the form $\mathbf{b} + (1/r)\mathbf{a}$, where $\mathbf{b} = (b_1, \dots, b_n)$ is a lattice point, which implies, by the definition of K' , that K contains the point $r\mathbf{b} + \mathbf{a}$. Thus for each one of the r^n specified choices of $\mathbf{a} = (a_1, \dots, a_n)$ there is a lattice point $\mathbf{c} = (c_1, \dots, c_n)$ in K such that $c_i \equiv a_i \pmod{r}$ for $i = 1, \dots, n$. Since the a_i satisfy $0 \leq a_i \leq r-1$, all these r^n lattice points \mathbf{c} in K are distinct, and therefore $N(K) \geq r^n$. This proves (3).

To prove (4) we utilize the well-known method of estimating the number of lattice points in a sphere. Let $S(x)$ be the n -dimensional unit sphere of radius x , centered at the origin. Then $V(S(\alpha n)) = \omega_n(\alpha n)^n$ and $A(S(\alpha n)) = n\omega_n(\alpha n)^{n-1}$, where

$$\omega_n = \frac{\pi^{n/2}}{\Gamma((n/2) + 1)}$$

is the volume of an n -dimensional unit sphere, and hence $V(S(\alpha n))/A(S(\alpha n)) = \alpha$. Now with each lattice point (a_1, \dots, a_n) belonging to the interior of $S(\alpha n)$ we associate the unit cube $\{(y_1, \dots, y_n) \mid a_i \leq y_i \leq a_i + 1\}$. The volume of the union of these cubes is $N(S(\alpha n))$ and they are all contained in a sphere of radius $\alpha n + \sqrt{n}$ with center at the origin (note that \sqrt{n} is the length of the main diagonal of the unit cube). Therefore

$$N(S(\alpha n)) \leq \omega_n(\alpha n + \sqrt{n})^n,$$

and this implies the upper estimate (4).

Although the lower estimate (3) is probably not the best possible, the bound (4) shows that as a function of α it is of the right order of magnitude. It seems likely, however, that (4) gives a better approximation to $l(\alpha, n)$ than (3); in fact, we conjecture that for each fixed $n \geq 1$ we have

$$l(\alpha, n) = \frac{\pi^{n/2}}{\Gamma((n/2) + 1)} n^n \alpha^n (1 + o(1)) \quad \text{as } \alpha \rightarrow \infty.$$

Note added in proof. For further results, including a proof of the above conjecture, see J. Bokowski and A. M. Odlyzko, *Lattice points and the volume 1 area ratio of convex bodies*, to appear in *Geometriae Dedicata*.

Acknowledgement. I would like to thank Professor T. M. Apostol for his help in the preparation of this paper.

References

1. E. A. Bender, Area-perimeter relations for two-dimensional lattices, this MONTHLY, 69 (1962) 742-744.

2. H. Hadwiger, Volumen und Oberfläche eines Eikörpers, der keine Gitterpunkte überdeckt, *Math. Zeit.*, 116 (1970) 191–196.
3. J. Hammer On a general area-perimeter relation for two-dimensional lattices, this *MONTHLY*, 71 (1964) 534–535.
4. ———, Some relatives of Minkowski's theorem for two-dimensional lattices, this *MONTHLY*, 73 (1966) 744–746.
5. J. M. Wills, Ein Satz über konvexe Mengen und Gitterpunkte, *Monatsh. Math.*, 72 (1968) 451–463.
6. ———, Ein Satz über konvexe Körper und Gitterpunkte, *Abh. Math. Sem. Univ. Hamburg*, (1970) 8–13.
7. ———, On lattice points and the volume/area ratio of convex bodies, this *MONTHLY*, 78 (1971) 47–49.

INCREASING CONTINUOUS SINGULAR FUNCTIONS

GERALD FREILICH, Queens College, (CUNY)

The Cantor function (sometimes referred to as the Lebesgue function) is the standard example of a monotone continuous singular function; though monotone, it is not strictly increasing. An ingenious example of a strictly increasing function with these properties is given in Riesz and Sz.-Nagy [2, pp. 48–49] and in Hewitt and Stromberg [1, pp. 278–282]. In this note, another method of constructing such functions is given.

We first review the definition of the Cantor function. Let S denote the Cantor ternary set defined by

$$S = \left\{ \sum_{n=1}^{\infty} 3^{-n} a_n : a_n = 0 \text{ or } 2 \text{ for all } n \right\}.$$

If $x = \sum_{n=1}^{\infty} 3^{-n} a_n$ and if for all n , $a_n = 0$ or 2 , then the Cantor function assigns to x the value $C(x) = \sum_{n=1}^{\infty} 2^{-n-1} a_n$. For $0 \leq y \leq 1$, the Cantor function assigns to y the value

$$C(y) = \sup\{C(x) : x \in S \text{ and } x \leq y\}.$$

We shall let C denote the usual Cantor function extended to the domain $(-\infty, \infty)$ by defining $C(x) = 0$ if $x \leq 0$ and $C(x) = 1$ if $x \geq 1$. It is well known that C is nondecreasing, continuous, and that $C' = 0$ almost everywhere.

We now give a simple construction of an increasing function with the above properties along with a simple proof of these properties.

CONSTRUCTION: Let $A = \{a_1, a_2, \dots\}$ be any countable set that is dense in the reals. Define

$$f(x) = \sum_{n=1}^{\infty} 2^{-n} C(2^n(x - a_n)).$$

ASSERTION. *The function f is continuous, strictly increasing, and singular.*

Proof. The uniform convergence of the sum defining f implies that f is continuous. If $x_2 < x_1$, choose a_n so that $x_2 < a_n < x_1$. Then

$$C(2^n(x_1 - a_n)) > 0 = C(2^n(x_2 - a_n)) \text{ and } C(2^m(x_1 - a_m)) \geq C(2^m(x_2 - a_m))$$

for $m \neq n$, so that $f(x_1) > f(x_2)$. By Fubini's theorem on differentiation of series with monotonic terms [2, pp. 11-12],

$$f'(x) = \sum C'(2^n(x - a_n)) = 0 \text{ for almost all } x.$$

References

1. E. Hewitt and K. Stromberg, *Real and Abstract Analysis*, Springer Verlag, New York, 1965, MR 32 # 5826.
2. F. Riesz and B. Sz.-Nagy, *Functional Analysis*, Ungar, New York, 1955, MR 17, 175.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

QUESTIONS ON A SEQUENCE OF ULAM

BERNARDO RECAMÁN, Colegio San Carlos, Bogotá, Colombia

Ulam [1] introduced sequences of positive integers constructed in the following way: "Let u_1 and u_2 be given integers; we construct an increasing sequence of integers by adjoining those which can be represented in just one way as the sum of two distinct preceding members of the sequence." I am concerned with the case $u_1 = 1$, $u_2 = 2$ and shall call members of this sequence ***U-numbers***. That there are infinitely many *U-numbers*, may be seen by supposing that they form a finite set; then the sum of the two largest would be a *U-number* not in the set, contradicting our supposition. The first few *U-numbers* are

1, 2, 3, 4, 6, 8, 11, 13, 16, 18, 26, 28, 36, 38, 47, 48, 53, 57, 62, 69, 72, 77, 82, 87, 97, 99, ...

From the many questions that can be asked about *U-numbers*, I choose the following:

1. Apart from $1 + 2 = 3$, can the sum of two consecutive *U-numbers* be a *U-number*? It cannot be the *next U-number*, as R. B. Eggleton (written communi-

cation) observes, since if $u_n + u_{n-1} = u_{n+1}$ is a unique sum of U -numbers, so is $u_n + u_{n-2}$, which lies between u_n and $u_n + u_{n-1}$.

2. Are there infinitely many numbers (such as 23, 25, 33, 35, 43, 45, 67, 92, 94, 96 ...) which are *not* the sum of two U -numbers?
3. (The problem with which Ulam introduced the sequences.) Do the U -numbers have positive density? In other words, if $U(x)$ is the number of U -numbers less than x , is $\liminf U(x)/x > 0$? The argument given above shows that $U(x) > \log_2 x$ since each U -number is less than twice the preceding; in fact the observation made under question 1 implies that $u_{n+1} \leq u_n + u_{n-2}$ and improves the estimate to $U(x) > \log_\alpha x$ where α is the real root of $x^3 - x^2 - 1 = 0$.
4. Are there infinitely many instances in which two consecutive integers are both U -numbers, such as 47 and 48? Clearly, three consecutive integers greater than $u_2 = 2$ cannot all be U -numbers.
5. Are there arbitrarily large gaps in the sequence of U -numbers?

Reference

1. S. M. Ulam, *Problems in Modern Mathematics*, Interscience, New York, 1964, p. ix.

EQUITABLE COLORING

WALTER MEYER, Adelphi University

Recently the notion of vertex colorings for graphs has proved useful in an operations research context [2]. In the application in question, vertices represented garbage collection routes and two such vertices were joined when the corresponding routes should not be run on the same day. The problem of assigning one of the six days of the work week to each route becomes the problem of six-coloring the graph. On practical grounds it might also be desirable to have an approximately equal number of routes run on each of the six days. This suggests:

DEFINITION. Suppose the vertices of a graph G are colored with p colors so that no edge joins vertices of the same color and the cardinalities of the color sets differ by at most one. Then G is said to be equitably p -colored. If p is the least integer n for which G can be equitably n -colored, then p is the **equitable chromatic number** of G , denoted by $\chi_e(G)$.

For example, for the 'star' graph $K_{1,n}$ in which one vertex is joined to each of n others, $\chi_e(K_{1,n}) = 1 + \{n/2\}$, where $\{x\}$ denotes the least integer not less than x .

The usual coloring methods do not seem to go far in studying equitable coloring, but we offer the following result.

THEOREM. *If T is a tree with maximum valence $\Delta(T)$, then T can be equitably colored with $\{\Delta(T)/2\} + 1$ colors.*

Proof. The proof is by induction on k , the number of vertices. The least value of k is $\Delta(T) + 1$ and this occurs only for $K_{1,n}$. By our earlier observation, the base of the induction is secure.

If $k > \Delta(T) + 1$ then it can be shown there will exist a vertex v with these properties:

1. $\text{val}(v) \geq 2$,
2. all but one, say x , of the vertices adjacent to v are 1-valent,
3. there exists a vertex v' of T , $v' \neq v$, where $\text{val}(v') = \Delta(T)$.

We shall denote $\text{val}(v)$ by t . Thus, the vertices other than x which are adjacent to v are x_1, x_2, \dots, x_{t-1} (Figure 1).

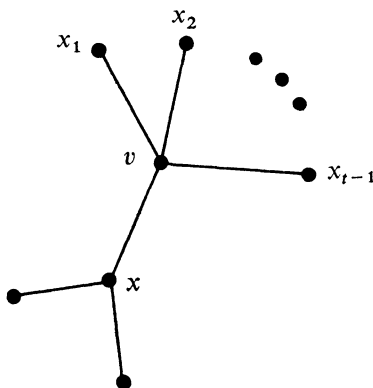


FIG.. 1

We prune the tree T by removing edge xv , thereby detaching a star with center v and leaving a tree T' which can be equitably colored with $\{\Delta(T)/2\} + 1$ colors by the inductive hypothesis.

Before proceeding further, it is convenient to introduce the color balance vector for a coloring of a graph with c colors. If w is the least cardinality of a color class in a c -coloring of a graph G , then the color balance vector of the coloring is (b_1, b_2, \dots, b_c) where b_i is the cardinality of the i th color set minus w . Thus a necessary and sufficient condition for a coloring to be equitable is that its color balance vector has only 0 and 1 as entries.

Now as we restore the pruned star to T' and recreate T , it is convenient to keep track of the color balance vector as we proceed to color the restored vertices. There are various cases according to the color balance of the coloring of T' . With no loss of generality one can assume in all cases that x is colored with the second color. The style of reasoning is similar in all cases, so we prove only the third case.

CASE 1. $b_i = 0$ for all $i \neq 2$.

CASE 2. There is an $i \neq 2$, where $b_i \neq 0$ and, in addition, the number of zero b_i , denoted z , is greater than or equal to t .

CASE 3. *There is an $i \neq 2$ ($i = 1$ say), where $b_i \neq 0$. Also $z \leq t - 1$.*

In case 3, color v with the first color and then use each color i for which $b_i = 0$ once to color one of the x_i . This produces the color balance vector $(1, 0, 0, \dots, 0)$. However, unless $z = t - 1$ not all the x_i have been colored. If uncolored vertices remain, we can run through the colors $2, 3, \dots, \{\Delta(T)/2\} + 1$ twice more without producing an inequitable color balance. Clearly this suffices to color each remaining x_i , since the number of these is $< t$ and $t \leq \Delta(T) \leq 2\{\Delta(T)/2\}$.

This theorem shows that, in general, $\chi_e(G)$ is not directly related to $\chi(G)$. A bound for $\chi_e(G)$ involving $\Delta(G)$ seems more reasonable. We offer the following conjecture whose validity has been verified for all graphs with 6 or fewer vertices.

CONJECTURE. $\chi_e(G) \leq \Delta(G)$ for all connected graphs G except the complete graphs and the odd circuits.

A reasonable place to begin on this conjecture might be with graphs where $\Delta(G) = 3$, for then a spanning 3-tree might already require 3 colors for an equitable coloring. Also, from the form of the conjecture, one is led to seek a connection with Brooks' Theorem [1].

Editorial Note: It follows from a result of Hajnal-Szemerédi [3] that $\chi_e(G) \leq \Delta(G) + 1$.

References

1. R. L. Brooks, On colouring the nodes of a network, Proc. Cambridge Philos. Soc., 37 (1941) 194–197; MR 6–281.
2. A. C. Tucker. Perfect graphs and an application to optimizing municipal services, SIAM-Review, 15 (1973).
3. A. Hajnal and E. Szemerédi, Proof of a conjecture of P. Erdős, in Combinatorial Theory and its Applications, Colloq. Math. Soc. János Bolyai, 4, (1970) 601–623.

CLASSROOM NOTES

EDITED BY ROBERT GILMER

Material for this Department should be sent to David Roselle, Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803.

THE LAGRANGE MULTIPLIER RULE

E. J. MCSHANE, University of Virginia

Because of its many uses, the Lagrange multiplier rule is presented in many advanced calculus texts. But the proof is customarily based on the implicit functions theorem, which to many undergraduates appears sophisticated. Moreover, as L. C.

Young remarks, the theorem for extrema subject to inequality constraints should be in the advanced calculus texts of the future. We here present a proof of the latter theorem that uses, beyond quite elementary calculus, only the theorem (Bolzano-Weierstrass) that from a bounded sequence of points in real n -space R^n one can always extract a convergent subsequence, and the theorem that a real function f defined and continuous on a closed ball $\{x \in R^n: |x - a| \leq r\}$ in R^n attains its minimum at some point in that ball. Even these two theorems are proved, e.g., in Bers' *Calculus*, pp. S4 and S6. The simpler proof of the theorem, as usually stated, is obtained by removing all references to inequality constraints.

The coordinates of a point x in R^n will be denoted by $x^{(1)}, \dots, x^{(n)}$. If f is defined on a set G in R^n and $a \in G$, the partial derivative at a of $(f(x): x \in G)$ with respect to the j th coordinate will be denoted by $D_j f(a)$, if it exists. Also, as usual,

$$\nabla f(a) = (D_1 f(a), \dots, D_n f(a)),$$

and

$$f^+(x) = \max(f(x), 0).$$

THEOREM. Let $f, g_1, \dots, g_k, h_1, \dots, h_s$ be defined and continuous, together with their partial derivatives $D_j f, D_j g_1, \dots, D_j g_k, D_j h_1, \dots, D_j h_s$ ($j = 1, \dots, n$), on a set G in R^n . Let x_0 be an interior point of G at which (i) the conditions

$$(1) \quad g_i(x) = 0 \quad (i = 1, \dots, k); \quad h_r(x) \leq 0 \quad (r = 1, \dots, s)$$

are satisfied, and (ii) $f(x_0) \leq f(x)$ for all x in G that satisfy (1). Then there are numbers $\lambda_0, \lambda_1, \dots, \lambda_k, \mu_0, \dots, \mu_s$ not all 0 such that

$$(2) \quad \lambda_0 D_j f(x_0) + \sum_{i=1}^k \lambda_i D_j g_i(x_0) + \sum_{r=1}^s \mu_r D_j h_r(x_0) = 0, \quad (j = 1, \dots, n).$$

Moreover,

- (i) $\lambda_0 \geq 0$ and $\mu_r \geq 0$ ($r = 1, \dots, s$);
- (ii) for each r such that $h_r(x_0) < 0$, $\mu_r = 0$;
- (iii) if the g_i and h_r satisfy the "Kuhn-Tucker constraint condition" that the gradients at x_0 of the g_i and those h_r for which $h_r(x_0) = 0$ are linearly independent vectors, it is possible to choose $\lambda_0 = 1$.

Without loss of generality we assume that $x_0 = (0, 0, \dots, 0)$; that $f(x_0) = 0$; and that

$$h_1(x_0) = \dots = h_z(x_0) = 0, \quad h_r(x_0) < 0 \quad (r = z+1, \dots, s).$$

We choose a positive ε_1 such that the closed ball $B(\varepsilon_1) = \{x \in R^n: |x| \leq \varepsilon_1\}$ is contained in G and $h_{z+1}(x), \dots, h_s(x)$ are negative on $B(\varepsilon_1)$. We first prove

(3) To each ε such that $0 < \varepsilon \leq \varepsilon_1$ there corresponds an N such that

$$f(x) + |x|^2 + N \left\{ \sum_{i=1}^k g_i(x)^2 + \sum_{r=1}^z [h_r^+(x)]^2 \right\} > 0$$

for all x such that $|x| = \varepsilon$.

Suppose false. Then there exist numbers N_1, N_2, \dots tending to ∞ and points x_1, x_2, \dots with $|x_m| = \varepsilon$ such that for all m

$$(4) \quad f(x_m) + |x_m|^2 \leq -N_m \left\{ \sum_{i=1}^k g_i(x_m)^2 + \sum_{r=1}^z h_r^+(x_m)^2 \right\}.$$

A subsequence of the x_m converges to a point x^* ; without loss of generality we may suppose this subsequence to be the whole sequence. Then $|x^*| = \lim |x_m| = \varepsilon$ and $f(x_m) \rightarrow f(x^*)$. By dividing both members of (3) by $-N_m$ and letting $m \rightarrow \infty$ we obtain

$$\sum_{i=1}^k g_i(x^*)^2 + \sum_{r=1}^z h_r^+(x^*)^2 = 0.$$

Therefore x^* satisfies (1), so $\lim f(x_m) = f(x^*) \geq f(x_0) = 0$. But by (4) $f(x_m) \leq -\varepsilon^2$. This contradiction establishes (3).

We next prove

(5) To each ε such that $0 < \varepsilon \leq \varepsilon_1$ there corresponds a point \bar{x} and a unit vector $(\lambda_0, \lambda_1, \dots, \lambda_k, \mu_1, \dots, \mu_z)$ with non-negative $\lambda_0, \mu_1, \dots, \mu_z$ such that $|\bar{x}| < \varepsilon$ and

$$\lambda_0 [D_j f(\bar{x}) + 2\bar{x}^{(j)}] + \sum_{i=1}^k \lambda_i D_j g_i(\bar{x}) + \sum_{r=1}^z \mu_r D_j h_r(\bar{x}) = 0 \quad (j = 1, \dots, n).$$

With the N of statement (3), we define F on G by

$$F(x) = f(x) + |x|^2 + N \left\{ \sum_{i=1}^k g_i(x)^2 + \sum_{r=1}^z h_r^+(x)^2 \right\}.$$

There is a point \bar{x} in the closed ball $B(\varepsilon)$ at which F assumes its least value on $B(\varepsilon)$; then $F(\bar{x}) \leq F(0) = 0$, so by (3) we cannot have $|\bar{x}| = \varepsilon$. So \bar{x} is interior to $B(\varepsilon)$, and all first-order partial derivatives of F must vanish at \bar{x} . At \bar{x} the function $[h_r^+]^2$ has derivatives

$$2h_r^+(\bar{x}) D_j h_r(\bar{x}),$$

obviously if $h_r^+(\bar{x}) > 0$ and because $[h_r^+]^2$ vanishes at least to second order where $h_r^+(x) = 0$. So for $j = 1, \dots, n$ we have

$$(6) \quad D_j f(\bar{x}) + 2\bar{x}^{(j)} + \sum_{i=1}^k 2N g_i(\bar{x}) D_j g_i(\bar{x}) + \sum_{r=1}^z 2N h_r^+(\bar{x}) D_j h_r(\bar{x}) = 0.$$

Define

$$\begin{aligned}
 L &= \left\{ 1 + \sum_{i=1}^k [2Ng_i(\bar{x})]^2 + \sum_{r=1}^z [2Nh_r^+(z)]^2 \right\}^{\frac{1}{2}}, \\
 \lambda_0 &= 1/L, \\
 \lambda_i &= 2Ng_i(\bar{x})/L \quad (i = 1, \dots, k), \\
 \mu_r &= 2Nh_r^+(\bar{x})/L \quad (r = 1, \dots, z), \\
 \mu_r &= 0 \quad (r = z + 1, \dots, s).
 \end{aligned}$$

Then $(\lambda_0, \lambda_1, \dots, \lambda_k, \mu_1, \dots, \mu_s)$ is a unit vector, and λ_0 and the μ_r are non-negative, and if we divide both members of (6) by L we obtain the equation in (5).

Now choose positive numbers $\varepsilon_1 > \varepsilon_2 > \varepsilon_3 > \dots$ tending to 0. For $m = 1, 2, \dots$ we choose a point \bar{x}_m with $|\bar{x}_m| < \varepsilon_m$ and a unit vector $(\lambda_{0,m}, \lambda_{1,m}, \dots, \mu_{z,m}, 0, \dots, 0)$ with non-negative $\lambda_{0,m}$ and $\mu_{r,m}$ such that the equation in (5) (with obvious notational changes) holds; this is possible, by (5). We choose a subsequence for which the unit vectors converge to a limit $(\lambda_0, \lambda_1, \dots, \lambda_k, \mu_1, \dots, \mu_s)$. Since $\bar{x}_m \rightarrow x_0$, the equation in (5) holds with this limit-vector and with x_0 in place of \bar{x} . The theorem is established except for conclusion (iii).

If the Kuhn-Tucker constraint condition holds we cannot have $\lambda_0 = 0$, since then (2) would contradict the linear independence of $\nabla g_1(x_0), \dots, \nabla g_k(x_0), \nabla h_1(x_0), \dots, \nabla h_z(x_0)$. So $\lambda_0 > 0$, and the multipliers $(1, \lambda_1/\lambda_0, \dots, \lambda_k/\lambda_0, \mu_1/\lambda_0, \dots, \mu_s/\lambda_0)$ satisfy all the requirements.

THE MINI-MAX PROPERTY OF THE TYCHONOFF PRODUCT TOPOLOGY

D. E. CAMERON, University of Akron

In 1923 Tietze [5] defined a topology for the Cartesian product of arbitrary topological spaces; this topology is often referred to as the box topology. The accepted definition for the product topology was given by Tychonoff [6] in 1930; it agrees with Tietze's topology on finite products but differs on infinite products. Whatever definition of the product topology is used, it is essential that for finite products it agree with the product topologies of both Tietze and Tychonoff because this definition is in keeping with the relation between the usual topology of the reals and the topology of Euclidean n -space. In this note we shall give a justification for the acceptance of the Tychonoff definition.

DEFINITION 1 (The Tychonoff Product Topology): If (X_α, T_α) for α in A are topological spaces, the product topology $\times\{T_\alpha: \alpha \in A\}$ on $\times\{X_\alpha: \alpha \in A\}$ —the Cartesian product of the X_α 's for α in A —is the topology which has as its subbase all sets of the form

$\prod_{\alpha}^{-1}(U_{\alpha})$, where U_{α} is in T_{α} for α in A and $\prod_{\alpha}: \times\{X_{\alpha}: \alpha \in A\} \rightarrow X_{\alpha}$ is the projection mapping for each α in A .

This definition differs from that of Tietze in that Tietze permitted arbitrary intersections of sets of the form $\prod_{\alpha}^{-1}(U_{\alpha})$ for his base while Tychonoff's definition permits only finite intersections of the same sets for its base.

DEFINITION 2. If T and T^* are two topologies on a nonempty set X , T is said to be *coarser* than T^* if $T \subset T^*$ (T^* is *finer* than T).

One sees by comparing the two definitions that Tietze's topology is finer than Tychonoff's on infinite products but that they are identical on finite products. Tychonoff's definition achieved its acceptability primarily because of the following result.

THEOREM 1. *The Tychonoff product topology is the coarsest topology on the Cartesian product for which all projection mappings are continuous [7, p 54–55].*

The argument to be considered here for the acceptance of the Tychonoff definition is that if the Tychonoff product topology were replaced by any other product topology then one of the following two theorems would be invalid.

THEOREM 2. *The Tychonoff product topology is a Hausdorff topology if and only if each coordinate space is a Hausdorff space [7, p. 87].*

THEOREM 3. *The Tychonoff product topology is a compact topology if and only if each coordinate space is a compact space [7, p. 120].*

The latter theorem is called the Tychonoff Product Theorem and is considered by many topologists to be the most important theorem in general topology. One reason for its importance is the fact that it is equivalent to the axiom of choice [3, p. 33 and p. 143–144; 4].

DEFINITION 3. Let R be a topological property, X a nonempty set. A topology T on X with property R is said to be *maximal R* (*minimal R*) if for any finer (coarser) topology T^* on X , T^* does not have property R .

Theorem 2 and Theorem 3 are related by the following theorem due to E. Hewitt [2, p. 320].

THEOREM 4. *A compact Hausdorff space is maximal compact and minimal Hausdorff.*

By using the product theorems previously stated, we see that the Tychonoff product of compact Hausdorff spaces is maximal compact and minimal Hausdorff. Therefore if the product topology used is coarser than the Tychonoff topology, products of Hausdorff spaces would not necessarily be Hausdorff. Also products of

metric spaces would not be metric spaces as can be seen by taking suitable products of $[0, 1]$ (the closed unit interval) with the usual topology which is a compact metric space and recalling the fact that all metric spaces are Hausdorff. Another property which would not be productive is T_1 -complete regularity since M. P. Berri [1] has shown that the minimal T_1 -completely regular topologies are precisely the compact Hausdorff topologies.

If the product topology used were finer than the Tychonoff topology (as in the case of the Tietze topology) then the products of compact spaces would not necessarily be compact. Therefore the Tychonoff product topology may be considered as being mini-max, a strong argument for its acceptance.

References

1. M. P. Berri, Minimal topological spaces, *Trans. Amer. Math. Soc.*, 108 (1963) 97–105.
2. E. Hewitt, A problem of set-theoretic topology, *Duke Math. J.*, 10 (1943) 309–333.
3. J. L. Kelley, *General Topology*, Van Nostrand, Princeton, N. J., 1955.
4. J. L. Kelley, The Tychonoff product theorem implies the axiom of choice, *Fund. Math.*, 37 (1950) 75–76.
5. H. Tietze, Über Analysis Situs, *Abhandl. Math. Sem. Univ. Hamburg*, 2 (1923) 27–70.
6. A. Tychonoff, Über die topologische Erweiterung von Räumen, *Math. Ann.*, 102 (1930) 544–561.
7. S. Willard, *General Topology*, Addison-Wesley, Reading, Mass., 1970.

ANOTHER APPROACH TO THE CUBIC INTERPOLATING SPLINE

B. H. ROSMAN, Framingham State College

1. Introduction. A topic that is permeating numerical analysis courses at all levels is the problem of constructing a cubic spline which interpolates to data values at the joints. In this note two common approaches to this problem are surveyed and a new variant to one of these approaches is discussed. A **cubic spline** $s(x)$ on $[a, b]$ with $n - 1$ interior joints $a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$ is a C^2 function such that on $[x_{i-1}, x_i]$, $s(x)$ is a cubic polynomial, i.e., $s(x)$ is a piecewise cubic with extra smoothness requirements at the joints.

2. The interpolation problem. The problem discussed here is: given data values f_i at the x_i , $i = 0, 1, 2, \dots, n$, find a cubic spline $s(x)$ (with joints at the x_i) such that $s(x_i) = f_i$; $i = 0, 1, \dots, n$. It follows easily from the above definition that any cubic spline has the form

$$(1) \quad s(x) = \sum_{i=0}^3 a_i x^i + \sum_{k=1}^{n-1} c_k (x - x_k)_+^3,$$

for some choice of real values a_i , c_k , where

$$(x - x_k)_+^3 = \begin{cases} 0, & x - x_k < 0 \\ (x - x_k)^3, & x - x_k \geq 0. \end{cases}$$

From (1), we see that this is a linear problem with $n + 1$ equations ($s(x_i) = f_i$) and $n + 3$ unknowns (a_i, c_k). To remove two degrees of freedom, end conditions are imposed. The most common conditions are (A): $s'(x_i) = f'_i$; $i = 0, n$ or (B): $s''(x_i) = 0$; $i = 0, n$, where f'_i is $f'(x_i)$, for the function to be fitted, $f(x)$, or, if we are fitting discrete values, f'_i is just another data value.

3. Methods of solution. For computational reasons, the system (1) is usually not solved directly. Often another linear system is derived which gives the spline *locally*, i.e., the correct cubic is derived for each cell $[x_{i-1}, x_i]$. The usual approaches for doing this differ according to whether end conditions (A) or (B) are given.

For condition (A) the approach is as follows (see [2, chap. 1], [3, chap. 4] for details). Using Hermite interpolation formulas a cubic $p_{i-1}(x)$ is derived in $[x_{i-1}, x_i]$ which takes on the values $p_{i-1}(x_j) = f_j$; and $p'_{i-1}(x_j) = s'_j$; $j = i - 1, i$. Similarly, in $[x_i, x_{i+1}]$, $p_i(x)$ is derived such that $p_i(x_j) = f_j$; $p'_i(x_j) = s'_j$, $j = i, i + 1$. Since we have imposed four conditions in each cell, this means that once we have determined the s'_j values, we know (by substituting into each local Hermite cubic) the spline in each cell. To get an easily solvable system for the s'_j values, we compute $p''_{i-1}(x_i)$ and $p''_i(x_i)$ and equate the results (since we impose C^2 smoothness). Doing this for $i = 1, 2, \dots, n - 1$, we get a tridiagonal strictly diagonally dominant linear system with conditions (A) used to eliminate the unknowns s'_0, s'_n ([2], [3]).

When conditions (B) are imposed, a different method of derivation is usually given (see [1, chap. 2], [4, chap. 1]). In each cell $[x_{i-1}, x_i]$, $s''(x)$ is a straight line of the form

$$(2) \quad s''(x) = s''_{i-1} + \frac{x - x_i}{h}(s''_i - s''_{i-1}),$$

where s''_i, s''_{i-1} are (as yet unknown) values of $s''(x)$ at x_i, x_{i-1} and $h = x_i - x_{i-1}$. To get from (2) to a local formula for $s(x)$, two integrations are performed, more constants are introduced (the s'_i and f_i values) and by algebraic manipulation, the s'_i values are eliminated and a tridiagonal strictly diagonally dominant system is obtained for the s''_i values with conditions (B) used to eliminate s''_0, s''_n . When this is solved $s(x)$ is then known in each cell.

4. Another approach. Although the two approaches lead to similar linear systems the method used for conditions (B) is more cumbersome. Some of my students and I have wondered why the more direct approach for conditions (A) cannot be used for conditions (B). The answer is it can, and the solution furnishes an instructive example of interpolation at work.

The idea is to find a cubic $p(x)$ such that $p(x_i) = f_i$; $p''(x_i) = f''_i$, $i = 0, 1$, where f_i, f''_i are arbitrary values. Then the approach for condition (A) can be imitated

leading directly to the linear system associated with conditions (B). (Here, we equate the *first* derivatives at x_i .) The interpolation problem posed here is a special case of a Hermite-Birkhoff interpolation problem, i.e., there may be gaps in the data conditions (the required derivatives of $p(x)$ at a given point are not consecutive) and as such may not possess a solution. (Specifically, consider n ordered pairs (i, j) , where i, j are integers with $1 \leq i \leq k \leq n$, $0 \leq j \leq n-1$. Let $x_1 < x_2 < \cdots < x_k$ be any set of numbers and for each of the (i, j) , let y_i^j be a given data value. The Hermite-Birkhoff interpolation problem is: find a polynomial $p(x)$ of degree at most $n-1$ such that $p^{(j)}(x_i) = y_i^j$ for each (i, j) .) However, here we do have a unique solution for by explicit computation the value of the determinant of the system for the coefficients a_i of the cubic $p(x) = \sum_{i=0}^3 a_i x^i$ has the value $-12(x_1 - x_0)^2$. (The determinant is derived from the conditions $p(x_i) = f_i$; $p''(x_i) = f_i''$ at the distinct points x_0, x_1 .)

The next step is to get a Lagrange type formula for the interpolating cubic. To this end we want cubics p_0, p_1, q_0, q_1 , such that $p_0''(x) = p_1''(x) = 0$ at $x = x_0, x_1$; $p_i(x_j) = \delta_{ij}$, $j = 0, 1$; and $q_0(x) = q_1(x) = 0$ at $x = x_0, x_1$, $q_i''(x_j) = \delta_{ij}$, $j = 0, 1$, where δ_{ij} is the Kronecker delta. For then we can write $p(x)$ as

$$(3) \quad p(x) = \sum_{i=0}^1 f_i p_i(x) + \sum_{i=0}^1 f_i'' q_i(x).$$

To get p_0, p_1 , note that p_0'' is a linear polynomial with two zeroes; hence $p_0'' \equiv 0$ so that $p_0(x)$ is just a straight line passing through the points $(x_0, 1), (x_1, 0)$. The same reasoning applies to p_1 , so that

$$(4) \quad p_0(x) = \frac{x - x_1}{x_0 - x_1}; \quad p_1(x) = \frac{x - x_0}{x_1 - x_0}.$$

To get q_0, q_1 , note that q_0 has zeroes at x_0, x_1 ; so we get that $q_0(x) = (x - x_0)(x - x_1)(ax + b)$, where a, b are picked to satisfy the conditions on $q_0''(x)$. Reasoning similarly for $q_1(x)$ we get

$$(5) \quad q_0(x) = \frac{\pi(x)(x + x_0 - 2x_1)}{6\pi'(x_0)}; \quad q_1(x) = \frac{\pi(x)(x + x_1 - 2x_0)}{6\pi'(x_1)},$$

where $\pi(x) = (x - x_0)(x - x_1)$.

Finally we remark that unfortunately this approach cannot be extended to higher degrees. For example consider the quintic $p(x)$ such that $p(x_i) = f_i$; $p''(x_i) = f_i''$, $i = 0, 1, 2$. With $x_0 = -1$; $x_1 = 0$; $x_2 = 1$, the determinant of the coefficient matrix for $p(x)$ is zero.

References

1. J. H. Ahlberg, E. N. Nilson, and J. L. Walsh, *The Theory of Splines and Their Applications*, Academic Press, New York, 1967.

2. T. N. E. Greville, *Theory and Applications of Spline Functions*, Academic Press, New York, 1969.
 3. T. J. Rivlin, *An Introduction to the Approximation of Functions*, Blaisdell, Waltham, Mass., 1969.
 4. B. Wendroff, *Theoretical Numerical Analysis*, Academic Press, New York, 1966.
-

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to Shirley Hill, Department of Mathematics, University of Missouri, Kansas City, MO 64110, or to Paul Mielke, Department of Mathematics, Wabash College, Crawfordsville, IN 47933.

ON BEHAVIORAL OBJECTIVES IN MATHEMATICS EDUCATION

L. C. JANSSON and R. T. HEIMER, Pennsylvania State University

In a recent issue on the MONTHLY, Allendoerfer [1] discussed with some insight the problem of the selection of elementary mathematics textbooks. He notes, correctly, we believe, that textbooks are marketed on the creative writing abilities of advertisers rather than on proven quantitative claims as to how well they teach or what their effects on students are.

Allendoerfer goes on to discuss the notion of "behavioral objectives" as a key to the writing and evaluation of textbooks. Though the term is slandered by referring to it as "psychological jargon," it appears to be of enough value that it serves as the basis for his entire article.

A behavioral objective is simply a very explicit statement of what it is one expects a student to be able to do following a unit of instruction. One commonly employed form contains three components:

- (1) *The Given*: the circumstances under which a particular performance is to take place. In a problem situation the "givens" of the problem would be included.
- (2) *The Required Performance*: an explicit statement of the task to be performed.
- (3) *The Criterion*: a statement of the conditions under which the performed task is judged to be satisfactory.

Thus we must state not only the desired goal behavior, but also the conditions under which it is to take place and the conditions of acceptable performance. An example of an instructional objective for a modern algebra course written in this form follows:

Recommendations. Even though the exceptions total only 7% it is our personal belief the objective test did a more realistic job of evaluating performance in the course. We feel that at least the 30 students in the unacceptable range would have been evaluated unfairly if multiple-choice tests had been used exclusively in determining grades. We would have preferred to discuss the differences in the test scores with certain students, especially those in the unacceptable range, but this proved to be impractical.

In addition to the possible incorrect evaluation factor, there is another equally important implication in using multiple-choice tests exclusively. While they can be effectively used to check mathematics skills we feel rather strongly that their exclusive use discourages the development of the student's ability to communicate using mathematics. Of course, this has not been verified statistically in this report, but our experiences with similar students in follow-up courses leads us to this belief. Certainly conscientious marking of homework papers can also help in the development of adequate mathematical communication skills. In conclusion, we suggest that college instructors make a thoughtful evaluation of their use of multiple-choice exams. We would be pleased to hear from others with thoughts on this subject.

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

ASSOCIATE EDITORS: JOSHUA BARLAZ, ERIC S. LANGFORD. COLLABORATING EDITORS: LEONARD CARLITZ, GULBANK D. CHAKERIAN, HASKELL COHEN, S. ASHBY FOOTE, ISRAEL N. HERSTEIN, MURRAY S. KLAMKIN, DANIEL J. KLEITMAN, ROGER C. LYNDON, MARVIN MARCUS, CHRISTOPH NEUGEBAUER, ALBERT WILANSKY, AND UNIVERSITY OF MAINE PROBLEMS GROUP: EARL M. L. BEARD, GEORGE S. CUNNINGHAM, CLAYTON W. DODGE, OSKAR FEICHTINGER, WILLIAM R. GEIGER, RAMESH GUPTA, GARY HAGGARD, PHILIP M. LOCKE, JOHN C. MAIRHUBER, CURTIS S. MORSE, GRATTAN P. MURPHY, EDWARD S. NORTHAM AND WILLIAM L. SOULE, JR.

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before January 31, 1974.

An asterisk () means neither the proposer nor the editors supplied a solution.*

E 2432*. *Proposed by L.-S. Hahn, University of New Mexico*

For each natural number n , let $f(n)$ be the smallest natural number that cannot be expressed as a combination of n n 's using only the rational operations of addition, subtraction, multiplication and division, together with unlimited use of parentheses. (No exponents, factorials, decimal points, etc., are allowed.) For example, $f(4) = 10$. Prove or disprove that $f(n)$ is an increasing function of n for $n \geq 3$.

E 2433. *Proposed by R. T. Smythe, University of Washington*

Let x_1, x_2, \dots be a sequence of real numbers. What is sometimes called Kronecker's Lemma asserts that if $\sum_{n=1}^{\infty} n^{-1} x_n$ is convergent, then $\bar{x}_n \rightarrow 0$ as $n \rightarrow \infty$, where $\bar{x}_n = n^{-1}(x_1 + \dots + x_n)$. (That is, $x_n \rightarrow 0$ in Cesàro mean.) The converse of this lemma is false: let $x_n = (\log(n+1))^{-1}$.

Prove the following partial converse: If $\bar{x}_n \rightarrow 0$ as $n \rightarrow \infty$, then $\sum_{n=1}^{\infty} n^{-1-\delta} x_n$ converges for all $\delta > 0$.

E 2434. *Proposed by George O'Brien, York University, Ontario*

Suppose that $\{a_n\}$ and $\{b_n\}$ are sequences of nonnegative numbers such that $(a_n)^n \rightarrow a$ and $(b_n)^n \rightarrow b$. Let p and q be nonnegative numbers such that $p + q = 1$: Evaluate $\lim(p a_n + q b_n)$. Generalize.

E 2435. *Proposed by Wells Johnson, Bowdoin College*

Let $p \geq 5$ be a prime number. Prove that there exist at least two distinct primes q_1, q_2 satisfying $1 < q_i < p-1$ and $(q_i)^{p-1} \not\equiv 1 \pmod{p^2}$.

E 2436. *Proposed by E. T. H. Wang, University of Waterloo*

An $n \times n$ matrix with nonnegative entries is *doubly stochastic* if each row sum and each column sum is 1. Characterize those doubly stochastic $n \times n$ matrices that commute with all doubly stochastic $n \times n$ matrices.

E 2437. *Proposed by David McLean, Warren, Michigan*

Let $\{a_n\}$ and $\{b_n\}$ be two sequences of real numbers that decrease monotonically to 0. If the series $\sum a_n$ and $\sum b_n$ both converge, then so does the series $\sum \max(a_n, b_n)$. Suppose that the series $\sum a_n$ and $\sum b_n$ both diverge. Does it follow that the series $\sum \min(a_n, b_n)$ must diverge?

SOLUTIONS OF ELEMENTARY PROBLEMS

$R[x]$ Can Contain \sqrt{x}

E 2370 [1972, 772]. *Proposed by John Hyde, Student, St. Olaf College*

Let R be a ring with identity, and let $R[x]$ be the ring of polynomials over R in

the indeterminate x . A current modern algebra textbook asks the student to prove that $R[x]$ cannot contain \sqrt{x} ; that is, $R[x]$ cannot contain a polynomial $f(x)$ such that $[f(x)]^2 = x$. Find an example of a ring R and a polynomial $f(x)$ that disproves this. Can R be commutative?

Solution by Robert Gilmer, Florida State University. For a ring R with identity and a positive integer n , let $R^{(n)}$ denote the ring of $n \times n$ matrices over R . If $n > 1$, then X has an n th root in $R^{(n)}[X]$. In fact, if A_n is the $n \times n$ matrix over R with 1 on the main subdiagonal and zeros elsewhere and if B_n is the $n \times n$ matrix over R with 1 in the upper right hand corner and zeros elsewhere, then it is easy to show that $(A_n + B_n X)^n = X$. The motivation for our choice of A_n and B_n comes from the canonical isomorphism ϕ between $R^{(n)}[X]$ and $(R[X])^{(n)}$; A_n and B_n are chosen so that $\phi(A_n + B_n X)$ is the companion matrix of $Y^n - X$ over $R[X]$. As a consequence, it follows that if S is the ring of infinite matrices over R , each of whose rows and columns contains only finitely many nonzero entries, then for each $k > 1$, X has a k th root in $S[X]$; if A is the element of S with A_k down the main diagonal and zeros elsewhere and if B is the element of S with B_k down the main diagonal and zeros elsewhere, then $(A + BX)^k = X$. More generally, if $r \in R$, then $(A + rB)^k = rI$, where I is the identity element of S . Thus if we identify R with its image under the isomorphism $r \rightarrow rI$ of R into S , then each element of R has a k th root in S , for each positive integer k .

If the ring R is commutative, then X has no n th root in $R[X]$ for $n > 1$, for if P is a proper prime ideal of R and if μ_P is the natural homomorphism from R onto R/P , then $\mu_P^*: R[X] \rightarrow (R/P)[X]$, defined by $\mu_P^*(\sum r_i X^i) = \sum \mu_P(r_i) X^i$, is a homomorphism such that $\mu_P^*(X) = X$, and since R/P is an integral domain, it is clear that X is not the n th power of an element of $(R/P)[X]$. It is of some interest to note that an element f of $R[X]$ may be an n th power modulo P for each proper prime ideal P of R while f is not an n th power in $R[X]$. This occurs, for example, if $R = \mathbb{Z}/(4)$ and $f = \bar{2} + X^n$.

Also solved by 43 others.

An Elusive Cubic Equation

E 2371 [1972, 772]. *Proposed by M. H. Greenblat*

A clever graduate student (CGS) was discussing a mathematical problem with his friend, the absent-minded professor (AMP). The CGS asked, "Do you remember the cubic equation we solved several weeks ago, you know the one in which the coefficients of all the terms were positive integers? It had integral roots, and the coefficient of the cubic term was unity?"

AMP — "Well, I remember it only vaguely."

CGS — "I'd like to reconstruct it. Do you remember the value of the constant term?"

AMP — “Not precisely. I remember it was either 2450 or 2540.”

CGS — “Well, do you remember the coefficient of the square term?”

AMP — “I’m afraid not, but it wouldn’t help you even if I did remember it.”
(In this, he underestimated the CGS.)

CGS — “Aha! Was the coefficient of the linear term as high as it could possibly be?”

AMP — “Yes.”

At this point, the CGS knew the equation in question. You can, too, with the above information.

Solution by the St. Olaf Problem Solving Group. Recall that

$$a = -\sum x_i, \quad b = \sum x_i x_j, \quad \text{and} \quad c = -x_1 x_2 x_3,$$

where x_1, x_2, x_3 are the roots of $x^3 + ax^2 + bx + c = 0$. Since a, b, c are positive, it follows that all roots are negative. Since $|x_1 x_2 x_3| = 2450$ or 2540 ($2 \cdot 5^2 \cdot 7^2$ or $2^2 \cdot 5 \cdot 127$), we list all possible combinations of three factors and check for cases which give identical linear coefficients. The only two are

$$x_1 = -5, \quad x_2 = -10, \quad x_3 = -49 \quad \text{and} \quad x_1 = -7, \quad x_2 = -7, \quad x_3 = -50.$$

The first case gives a linear coefficient of 785 and the second 749. Since the linear coefficient is to be as large as possible, we must have the polynomial equal to

$$(x + 5)(x + 10)(x + 49) = x^3 + 64x^2 + 785x + 2450.$$

Also solved by Anders Bager (Denmark), Larry Byrd & Truett Mathis, D. W. Erbach (England), and the proposer.

Editorial Note. No solver mentioned that no root could be positive by Descartes’ rule of signs. C. W. Karns stated that “the statement that AMP underestimated the ability of CGS implies that $x_1 + x_2 + x_3$ is unique.” He then finds all values for the coefficient b , the maximum value yielding $a = 2542$, $b = 5081$, and $c = 2540$ (using the notation of the above solution). Twenty-one other solvers arrived at the same solution by ignoring the comment about the square term coefficient.

Diagonals in a 0-1 Matrix

E 2372 [1972, 773]. *Proposed by E. T. H. Wang, University of Waterloo, Canada*

Let A be an $n \times n$ matrix with entries zero and one, such that each row and each column contains precisely k ones. A *generalized diagonal* of A is a set of n elements of A such that no two elements appear in the same row or the same column. Show that A has at least k pairwise disjoint generalized diagonals, each of which consists entirely of ones.

Solution by Mike Vitale, Skidmore College. The solution proceeds by induction. If $k = 1$, pick the n ones of A as the generalized diagonal. Suppose the assertion is true for $1, 2, \dots, k-1$, and each row and column of A contains exactly k ones. We first choose a single generalized diagonal of A which consists entirely of ones. Pick a one from each of rows $1, 2, \dots, k$ at random, subject only to the condition that no two are in the same column. This can always be done, since each row contains k ones. Suppose each of the ones in row $k+1$ lies in a column which has already been chosen. (If not, pick an available one from this row, and proceed.) In at least one of the rows $1, 2, \dots, k$ there is a one which does not lie above a one in row $k+1$. (For if not, the columns containing the ones of row $k+1$ have at least $k+1 > k$ ones in them, a contradiction.) Call this row i . Replace the element chosen from row i originally by the element not lying above a one in row $k+1$, and from row $k+1$ choose the one lying below the original element chosen from row i . Now go to row $k+2$; again, if there is an available one, choose it and go on to the next row. If not, there must be a row $1, 2, \dots, i-1, i+1, \dots, k+1$ which contains a one not lying above a one in row $k+2$. Perform the switch outlined above and go on to the next row.

In this way we can choose a generalized diagonal of A which consists entirely of ones. Replace each of the ones chosen by a zero to obtain a new matrix A' which has exactly $k-1$ ones in each row and column. By induction, we can find $k-1$ pairwise disjoint generalized diagonals which consist entirely of ones. Together with the first generalized diagonal chosen, these form a set of k generalized diagonals consisting entirely of ones.

Also solved by L. W. Beineke & R. E. Pippert, The Bennett College Team, R. B. Eggleton, Bill Knight, Joel Levy, L. E. Mattics, Henryk Minc, K. R. Rebman, Richard Sinkhorn, Phil Tracy, and the proposer.

Editor's Note. The solution above is self contained. Several readers pointed out that the result in question is an immediate consequence of well-known theorems in the literature. Minc cites D. König, *Theorie der endlichen und unendlichen Graphen* (1936), chap. XIV, s. 3, Proposition B; also M. Marcus and H. Minc, *Modern University Algebra* (1966), Theorem 4.3.

Balancing Integral Weights

E 2374 [1972, 905]. *Proposed by Judith Q. Longyear, Pennsylvania State University*

Suppose that $a_1 \leq a_2 \leq \dots \leq a_n$ are natural numbers such that $a_1 + \dots + a_n = 2n$ and such that $a_n \neq n+1$. Show that if n is even, then for some subset K of $\{1, 2, \dots, n\}$ it is true that $\sum_{i \in K} a_i = n$. Show that this is true also if n is odd when we make the additional assumption that $a_n \neq 2$.

Solution by J. G. Mauldon, Amherst College. Let $s_k = a_1 + \dots + a_k$ for $k = 1, 2, \dots, n-1$ and consider the set of $n+1$ (distinct) numbers

$$\{0, a_1 - a_n, s_1, s_2, \dots, s_{n-1}\}.$$

By the Pigeonhole Principle, at least one pair of elements are congruent to each other modulo n . We distinguish four cases.

(1) If $a_1 - a_n \equiv 0 \pmod{n}$, then $a_1 = a_n$ since $0 \leq a_1 - a_n \leq -n + 1$. Thus in this case it follows that $a_1 = \dots = a_n = 2$. If $n = 2m$ is even, then obviously any m -element subset K will do, whereas if n is odd, this circumstance is prohibited by assumption.

(2) If $s_i \equiv s_k \pmod{n}$ for some $1 \leq i < k \leq n - 1$, then since $2n - 2 \geq s_k - s_i \geq 1$ it follows that $s_k - s_i = n$, i.e., $a_{i+1} + \dots + a_k = n$, and we can take

$$K = \{i + 1, i + 2, \dots, k\}.$$

(3) If $s_k \equiv 0 \pmod{n}$ for some $1 \leq k \leq n - 1$, then $1 \leq s_k \leq 2n - 1$, so that necessarily $s_k = n$, and we can take $K = \{1, 2, \dots, k\}$.

(4) If $s_k \equiv a_1 - a_n \pmod{n}$ for some $1 \leq k \leq n - 1$, then either $k = 1$, implying $a_n \equiv 0 \pmod{n}$ and thus $a_n = n$ (since $a_n \leq n$) so that $K = \{n\}$ will do, or else $k \geq 2$ in which case $a_2 + \dots + a_k + a_n \equiv 0 \pmod{n}$. In this latter case, $1 \leq a_2 + \dots + a_k + a_n < a_1 + a_2 + \dots + a_k + \dots + a_n = 2n$ so that $a_2 + \dots + a_k + a_n = n$, and we can take $K = \{2, 3, \dots, k, n\}$.

Also solved by Bennett College Team, M. T. Bird, R. B. Eggleton, C. S. Gardner, S. I. Gendler, R. A. Gibbs & H. S. Stocker, Michael Goldberg, M. G. Greening (Australia), Erwin Just, Jonathan Kane, E. M. Klein, O. P. Lossers (Netherlands), John MacBain, W. D. Markel, W. W. Parsons, Kenneth Rosen, Michael Shimshoni (Israel), Alan Stein, Guy Torchinelli, and Charles Wexler.

Editorial Comment: This problem has the following interesting interpretation. Suppose that one has a number of blocks of integral weight, which average 2 weight units apiece, and which have the property that the heaviest block is not heavier than all the rest of the blocks put together. Then, with the exception of the case where one has an odd number of blocks all of weight 2, it is always possible to separate the blocks into two groups of equal weight. Moreover, it can be shown that the following algorithm will always lead to a solution of the problem: Take a two-pan balance, and starting with the heaviest weight and proceeding step-by-step with the next heaviest, etc., place the weights one at a time on that pan of the balance which is the lighter (if the pans are balanced, choose either). Then the placing of the last (lightest) block must necessarily make the two pans balance exactly.

Difference Sets of Commutative Groups

E 2375 [1972, 905]. Proposed by H. Kestelman, University College, London, England

Let G be an abelian group. For any subset S of G , let $D(S)$ denote the set of differences $x - y$, where $x, y \in S$. Show that if A and B are any subsets of G such that $G = A \cup B$, then either $D(A) \supseteq B$ or $D(B) \supseteq A$. Show further that if $G = A \cup B$ and if A and B are not disjoint, then $D(A) = G$ or $D(B) = G$.

Solution by John Coolidge, Florida State University. If $B \not\subseteq D(A)$, then there

exists an element $b \in B \setminus D(A)$. For any $a \in A$, $b + a \notin A$, since this would imply $b = (b + a) - a \in D(A)$; hence $b + a \in B$. Thus $a = (b + a) - b \in D(B)$ and $A \subseteq D(B)$.

Let $S - x = \{s - x \mid s \in S\}$. Since $A \cap B \neq \emptyset$, let $x \in A \cap B$. Obviously $0 \in (A - x) \cap (B - x)$ and $G = (A - x) \cup (B - x)$. By the preceding result assume $(A - x) \subseteq D(B - x)$. Since $0 \in B - x$, $B - x \subseteq D(B - x)$. Therefore $G \subseteq D(B - x)$. Also $D(B - x) \subseteq D(B)$ because $(b_1 - x) - (b_2 - x) = b_1 - b_2 \in D(B)$. Hence $G = D(B)$.

Also solved by S. Baskaran (India), Bennett College Team, D. M. Bloom, W. E. Boddien, Ramón Casanova, John Christopher, Alonzo Church, Jr., J. Coolidge, Patty DeLoach, Mary Ellien, Ali Fora (Jordan), Scott Forrest, C. S. Gardner, J. D. Gillam, M. G. Greening (Australia), J. W. Grossman, J. Haja, N. E. Harrison, C. V. Heuer, Erwin Just, Jonathan Kane, Charlotte Krauthamer, G. D. Kruse, S. E. Landsburg, O. P. Lossers (Netherlands), Luther College Problems Seminar, Milan Lustig (Czechoslovakia), D. E. Manes, L. E. Mattics & M. E. von Wolff, Gary McDonald & Merry McDonald, D. L. Milgram & Azriel Rosenfeld, C. L. Morgan, C. M. Price, Kenneth Rosen, W. J. Sánchez, George Schillinger, J. R. Smith, Christopher Steel, Temple University Problem Solving Group, Guy Torchinelli, Mike Vitale, Albert Wilansky, Qazi Zameeruddin (India), and the proposer.

Editor's Comment: Gary and Merry McDonald showed that the restriction that G be abelian can be omitted only if G is finite.

Prime Powers and the Sigma Function

E 2376 [1972, 906]. *Proposed by Arthur Marshall, Madison, Wisconsin*

Suppose that p and q are odd primes and that a and b are natural numbers such that $p^a > q^b$. Show that if p^a divides the product $\sigma(p^a)\sigma(q^b)$, then in fact $p^a = \sigma(q^b)$.

Solution by Temple University Problem Solving Group. We have

$$\sigma(p^a) = 1 + p + \cdots + p^a \equiv 1 \pmod{p}.$$

Therefore $(p^a, \sigma(p^a)) = 1$. Thus, if p^a divides $\sigma(p^a)\sigma(q^b)$, then $p^a \mid \sigma(q^b)$. Note that

$$\sigma(q^b) = 1 + q + \cdots + q^{b-1} + q^b = \frac{q^b - 1}{q - 1} + q^b < 2q^b.$$

But $q^b < p^a$. Therefore $\sigma(q^b) < 2p^a$. Also $1 < \sigma(q^b)$, so $p^a \mid \sigma(q^b)$ implies that $\sigma(q^b) = p^a$.

Also solved by Anders Bager (Denmark), Merrill Barnebey, Jay Boekhoff, T. B. Carroll, John Christopher, Scott Forrest, Joe Illick, Eleanor G. Jones, J. W. King, E. M. Klein, N. J. Kuenzi & Bob Prielipp, L. Kuipers, S. E. Landsburg, P. A. Lindstrom, O. P. Lossers (Netherlands), Andrzej Makowski (Poland), J. B. Muskat (Israel), Kenneth Rosen, T. Šalát (Czechoslovakia), Hwa Tang, Guy Torchinelli, Phil Tracy, W. R. Umbach (W. Germany), Mike Vitale, Roger Weitzenkamp, Charles Wexler, and the proposer.

Editorial Note. Several solvers pointed out that the primes need not necessarily be odd. The solution above is valid also if either p or q is 2.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers — The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before January 31, 1974. Contributors (in the United States) who desire acknowledgement of receipt of their solutions are asked to enclose self-addressed, stamped postcards.

5928. *Proposed by Carl Pomerance*

If n is an integer greater than 1, let $t(n)$ equal the number of squares in the ring $\mathbb{Z}/(n)\mathbb{Z}$. Find a formula for $s(n) = \lim_{k \rightarrow \infty} t(n^k)n^{-k}$.

5929. *Proposed by Frederick Stern, California State University at San Jose*

Let $0 \leq x < 1$. Show that

$$\left(\prod_{k=1}^{\infty} (1 - x^k) \right)^{-1} = 1 + \sum_{m=1}^{\infty} \sum_S \frac{x^{i_1 + i_2 + \dots + i_m}}{(1 - x^{i_1})(1 - x^{i_2}) \dots (1 - x^{i_m})},$$

where $S = (i_1, i_2, \dots, i_m)$ is any set of positive integers such that $i_1 < i_2 < \dots < i_m$.

5930. *Proposed by J. J. Buckley, University of South Carolina*

Let \mathcal{A} be the Lebesgue subsets of R and let \mathcal{B} be the Borel subsets of R . A problem in Halmos, *Measure Theory*, p. 143 implies that if $f: R \rightarrow R$ is Lebesgue measurable, then the graph of f belongs to the product σ -algebra $\mathcal{A} \times \mathcal{B}$. Is the converse true?

5931. *Proposed by C. J. Smyth, University of Turku, Finland*

Let γ be a PV-number (an algebraic integer with $|\gamma| > 1$, and conjugates $\gamma = \gamma_1, \gamma_2, \dots, \gamma_n$, with $|\gamma_i| < 1$ ($i = 2, \dots, n$)). Show that for $i \neq j$, $|\gamma_i| = |\gamma_j|$ implies $\gamma_i = \bar{\gamma}_j$.

5932. *Proposed by Richard Stanley, University of California at Berkeley*

Call two permutations σ and ρ in the symmetric group S_n equivalent (denoted $\sigma \sim \rho$) if every cycle c in the disjoint cycle decomposition of σ is some power (depending on c) of a cycle in the disjoint cycle decomposition of ρ . It is easily seen that \sim is an equivalence relation. Let $E(n)$ denote the number of equivalence classes. Show that

$$\sum_{n=0}^{\infty} E(n) x^n / n! = \exp \left(\sum_{i=1}^{\infty} x^i / i \phi(i) \right),$$

where ϕ is the Euler totient function.

5933. *Proposed by P. L. Renz, Wellesley College*

Let \mathcal{G} be a family of random graphs constructed on a fixed countably infinite

vertex set V and having the property that the probability of $\{v, v'\}$ being an edge of a graph G of \mathcal{G} is p , with $0 \leq p \leq 1$, for each distinct pair of vertices v and v' in V . What is the probability that a graph G in \mathcal{G} contains an infinite complete subgraph?

SOLUTIONS OF ADVANCED PROBLEMS

k -Chromatic Graphs

5855 [1972, 523]. *Proposed by Ioan Tomescu, Ploiesti, Rumania*

Show that any k -chromatic graph on n vertices none of which are isolated must have at least $\frac{1}{2}k(k-1) + \frac{1}{2}(n-k)$ edges.

Solution by A. M. Hobbs, Texas A and M University. Choose a k -coloring of a k -chromatic graph G . For each color c , if no vertex of color c is adjacent to vertices of all other colors, every vertex of color c can be recolored. Since this is impossible, for each of the k colors there is a vertex of degree $k-1$ or more which is colored with that color. We have a lower bound for the degree of only one vertex of each color; for each of the other $n-k$ vertices there is at least one edge incident to it (no isolated vertices) and so the sum of the degrees in G is $\geq k(k-1) + (n-k)$. Thus the number of edges in G equals half of the sum of the degrees of the vertices in G , i.e., is $\geq \frac{1}{2}[k(k-1) + (n-k)]$.

Also solved by B. A. Broemser, Jacques Chone (France), D. J. Kleitman, O. P. Lossers (Netherlands), John Mitchen, B. R. Myers, Thomas Peterson, Simeon Reich (Israel), D. P. Sumner, Jeanne K. Tamaki, and the proposer.

Editorial Note. Broemser with his solution observes that the minimum number of edges is either attainable (n, k , same parity) or is attained by adding $1/2$ (n, k opposite parity).

Identities for Finite Subsets of a Set

5856 [1972, 523]. *Proposed by Jan Mycielski, University of Colorado*

For any collection X of finite subsets of a set S we denote by X^* the collection of all finite subsets T of S such that the number of subsets of T which belong to X is odd. Prove that $X^{**} = X$ and $(X\Delta Y)^* = X^*\Delta Y^*$, where $X\Delta Y = (X \cup Y) - (X \cap Y)$.

Solution, composite of contributions by Fred Galvin, University of California at Los Angeles, and R. W. Quackenbush, University of Manitoba. Let $F(S)$ denote the collection of all finite subsets of S . Note that

$$X^* = \{T \in F(S) : |2^T \cap X| \text{ is odd}\}.$$

$$1. \quad |X \cap 2^T| + |Y \cap 2^T| = |(X\Delta Y) \cap 2^T| + 2|(X \cap Y) \cap 2^T|,$$

whence

$$|(X\Delta Y) \cap 2^T| = |X \cap 2^T| + |Y \cap 2^T| \pmod{2},$$

and thus $(X\Delta Y)^* = X^*\Delta Y^*$.

2. $\emptyset^* = \emptyset$, so $\emptyset^{**} = \emptyset$.
3. If $X = \{T\}$, then $X^* = \{R \in F(S) : T \subseteq R\}$, and $X^{**} = X$.
4. For finite X , $X^{**} = X$ is proved by induction using (1), (2) and (3).
5. Let $T \subseteq S$. For $X \subset F(T)$, define

$$X' = \{U \in F(T) : |2^U \cap X| \text{ is odd}\}.$$

Then for $X \subset F(S)$, $U \in F(T)$, we will have $U \in (X \cap 2^T)'$ if and only if

$$|2^U \cap (X \cap 2^T)| = |X \cap 2^U|$$

is odd, if and only if $U \in X^* \cap 2^T$. Thus $(X \cap 2^T)' = X^* \cap 2^T$ for all $T \subseteq S$ and all $X \subset F(S)$.

Now let $X \subset F(S)$ be infinite and $T \in F(S)$. Then $T \in X^{**}$ if and only if

$$T \in X^{**} \cap 2^T = (X^* \cap 2^T)' = (X \cap 2^T)' = X \cap 2^T$$

(since T is finite) if and only if $T \in X$. Thus $X^{**} = X$.

Also solved by A. Ehrenfeucht, Ellen Hertz, Eitan Lapidot (Israel), William Massey, L. E. Mattics, P. L. Renz, and the proposer.

Note 1. Galvin suggests a generalization: Let P be a finite partially ordered set. If $A \subseteq P$, let $A^* = \{x \in P : |\{y \in A, y \leq x\}| \text{ is odd}\}$. Then

- (i) $(A \Delta B)^* = A^* \Delta B^*$ for all $A, B \subseteq P$.
- (ii) $A^{**} = A$ for all $A \subseteq P$, if and only if $|\{x \in P, a < x < b\}|$ is even for all $a, b \in P$.

Note 2. The proposer adds the following comment: Let $S = \{1, \dots, n\}$. It is well known that every Boolean function $f(x_1, \dots, x_n)$ can be represented in two ways

$$f(x_1, \dots, x_n) = \bigvee_{T \in X} \left(\bigwedge_{i \in T} x_i \wedge \bigwedge_{i \in S-T} \neg x_i \right),$$

$$f(x_1, \dots, x_n) = \sum_{T \in Y} \prod_{i \in T} x_i \pmod{2}$$

where X and Y are suitable families of subsets of S . The relationship between X and Y is precisely $X^* = Y$.

A Sum of (0, 1) Random Variables

5857 [1972, 523]. *Proposed by Gérard Letac, Institut Universitaire de Technologie, Aubière, France*

$X_1, X_2, \dots, X_t, \dots$ being independent random variables such that $P(X_i = 0) = P(X_i = 1) = \frac{1}{2}$, define $S_t = \sum_{i=1}^t X_i / 2^i$. Take a set H of rational numbers of the form $a/2^b$, such that H is dense in $[0, 1]$. Prove or disprove that $P(\exists t > 0 : S_t \in H) = 1$.

Solution by Ellen Hertz, Bronx, N.Y. Let $\varepsilon > 0$. Then there exists a set H of rationals of the form $a/2^b$ such that H is dense in the unit interval but $P(\exists t : S_t \in H) < \varepsilon$.

Proof. S_t is distributed uniformly on the 2^t points $0, 1/2^t, \dots, (2^t - 1)/2^t$. Let D be the set of all rationals in $(0, 1)$ of the form $a/2^n$. If $p \in D$ define $n(p)$ by setting $p = a/2^{n(p)}$, a odd. For any given $p \in D$, $P(S_t = p) = 1/2^t$ if $t \geq n(p)$ and $P(S_t = p) = 0$ if $t < n(p)$. Hence

$$P(\exists t: S_t = p) \leq \sum_{t=n(p)}^{\infty} 1/2^t = 1/2^{n(p)-1}.$$

Since there are only finitely many values of p for each value of $n(p)$, then $n(p)$ is unbounded in any subinterval of $(0, 1)$.

Construct H as follows: Let $1/2^{n_0-2} < \varepsilon$.

Step 1. Select $0 < p_1 < \frac{1}{2}$ such that $n(p_1) \geq n_0$. Select $\frac{1}{2} < p_2 < 1$ such that $n(p_2) > n(p_1)$.

Step n . Select values of $p_{n(n+1)/2+i}$, $i = 0, 1, \dots, n$ such that $i/(n+1) < p_{n(n+1)/2+i} < (i+1)/(n+1)$ always requiring that $n(p_{k+1}) > n(p_k)$, $k = 1, 2, \dots$.

Then

$$P(\exists t: S_t \in H) \leq \sum_{k=1}^{\infty} P(\exists t: S_t = p_k) \leq \sum_{k=1}^{\infty} 1/2^{n(p_k)-1} \leq \sum_{k=n_0-1}^{\infty} 1/2^k = 1/2^{n_0-2}.$$

Also solved by J. C. Kieffer, S. P. Lloyd, L. E. Mattics, Daniel Mosenkis, P. van der Steen (Netherlands), and the proposer.

An Irrational in a Covering of the Rationals

5858 [1972, 523]. *Proposed by Leonard Gallagher, University of Colorado*

Let $Q = \{r_i\}_{i=1}^{\infty}$ be any enumeration of the rationals and consider open intervals $I_i'' = N_{\frac{1}{2}n}(r_i)$ about r_i . Since

$$G = \bigcap_{n=1}^{\infty} \bigcup_{i=1}^{\infty} I_i^{n+h}$$

is a G_δ set, $Q \neq G$. Demonstrate an irrational element of the G_δ set.

Solution by G. A. Heuer, Concordia College. For each i , let $r_i = m_i/n_i$, where m_i and n_i are relatively prime integers and $n_i > 0$. Let $q_1 = r_1$ and $\phi(1) = 1$. When $q_k = r_{\phi(k)}$ is given, let $q_{k+1} = r_{\phi(k+1)}$, where $\phi(k+1)$ is the first integer greater than $\phi(k)$ such that $n_{k+1} > n_k$ and $|q_{k+1} - q_k| < \min\{1/2n_k^{n_k}, 1/2^{\phi(k)+k+1}\}$. Since $n_k \geq k$, the sequence $\{q_k\}$ converges to some number ξ .

For each h ,

$$\begin{aligned} |\xi - r_{\phi(h)}| &= |\xi - q_h| \leq \sum_{k=h}^{\infty} |q_{k+1} - q_k| \\ &< \min \left\{ \sum_{k=h}^{\infty} 1/2n_k^{n_k}, \sum_{k=h}^{\infty} 1/2^{\phi(k)+k+1} \right\} \end{aligned}$$

$$\begin{aligned}
&< \min \left\{ \sum_{k=h}^{\infty} 1/(2n_h^{n_h} \cdot 2^{k-h}), \sum_{k=h}^{\infty} 1/2^{\phi(h)+2k-h+1} \right\} \\
&\leq \min \{1/n_h^{n_h}, 1/2^{\phi(h)+h}\},
\end{aligned}$$

where in the next to the last step we use $(n_{k+1})^{n_{k+1}} > (n_k)^{n_{k+1}} \geq 2n_k^{n_k}$ and $\phi(k+1) \geq \phi(k) + 1$. Thus ξ is a Liouville number (hence transcendental) and $\xi \in I_{\phi(h)}^{\phi(h)+h}$ for each h . Therefore $\xi \in G$.

Also solved by R. J. Evans, O. P. Lossers (Netherlands), R. H. Marty, Lieselotte Miller, and the proposer.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN

with the assistance of the mathematics departments of St. Olaf and Carleton Colleges

COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER, CARLETON COLLEGE

We invite readers to submit reviews of significant recent college-level mathematics books. We especially encourage reviews based on classroom use, or comparative reviews of several related books. Reviews should ordinarily not exceed two pages (per book) typed double spaced. Manuscripts of reviews as well as books submitted for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057.

Set Theory and Metric Spaces. By Irving Kaplansky. Allyn and Bacon, Boston, Mass. 1972. xii + 140 pp. \$9.95. (Telegraphic Review, June/July 1972.)

How enjoyable it is to find a book in mathematics that is just plain fun to read. Unfortunately, this doesn't happen very often. It's a pleasure to report that *Set Theory and Metric Spaces* is one of these enjoyable additions to the field.

Irving Kaplansky returns in this case to speak to the undergraduate on what sounds to be a very ordinary topic. But to say "speak" is the main reason the book is not ordinary, for the book has the ring of the small classroom lecture with all of its freedom, humor, and patience. The author does this through an amazingly fresh and descriptive use of language — a conversation with the reader.

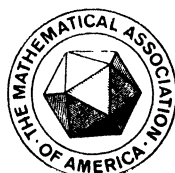
The book material itself has grown from a set of notes initially prepared by Spanier and used as a text at the University of Chicago. Kaplansky's goal is to present the working essentials of set theory that a mathematician needs to know along with some basic topics in the study of metric spaces.

In this philosophy, the author's introductory set theory is initially terse and somewhat intuitive. As he states himself, his set theory is "supernaive". He presents just

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF
THE MATHEMATICAL ASSOCIATION OF AMERICA



VOLUME 80

NUMBER 9

CODEN: AMMYAE

CONTENTS

The Spinor Spanner	E. D. BOLKER	977
Linear Combinations of Sets of Consecutive Integers	D. A. KLARNER AND R. RADO	985
The Equation $x'(t) = ax(t) + bx(t - \tau)$ with "Small" Delay	R. D. DRIVER, D. W. SASSER AND M. L. SLATER	990
The Quaternion Calculus	C. A. DEAVOURS	995
Inequalities for Sums of Distances	G. D. CHAKERIAN AND M. S. KLAMKIN	1009
The William Lowell Putnam Mathematical Competition	J. H. MCKAY	1017
Simple Groups		1028

MATHEMATICAL NOTES

On a Problem Concerning Euler's Phi-function	HAROLD DONNELLY	1029
What is the Probability that Two Group Elements Commute?	W. H. GUSTAFSON	1031
Remarks on the Bessel Polynomials	C. W. BARNES	1034
A Micronote on a Functional Equation	H. N. SHAPIRO	1041
An Addendum to the Paper "A Characterization of the $n \times n$ Matrices over a Finite Field"	J. V. BRAWLEY AND L. CARLITZ	1041

RESEARCH PROBLEMS

Exploring a Planet	L. FEJES TÓTH	1043
What are the Latin Square Groups?	J. J. CARROLL, G. A. FISHER, A. M. ODLYZKO, AND N. J. A. SLOANE	1045

CLASSROOM NOTES

Geometric Fit of a Monotonic Cubic	W. P. COOKE	1047
A Familiar Combinatorial Identity Proved by Complex Analysis	STEVEN MINSKER	1051

(Continued on inside cover)

NOVEMBER

1973

MATHEMATICAL EDUCATION

A Discovery Course in Graph Theory	J. L. LEONARD	1052
A Bow to Relevancy	R. L. WILSON	1053
Concerns of Two-year Colleges		
. G. F. GILMER, H. B. Siner, R. Mansfield and W. G. Chinn		1055
ELEMENTARY PROBLEMS AND SOLUTIONS		1057
ADVANCED PROBLEMS AND SOLUTIONS		1067
REVIEWS		1072
NEWS AND NOTICES		1090
MATHEMATICAL ASSOCIATION OF AMERICA		1090
Employment Information for Mathematicians		1090
April Meeting of the Indiana Section		1091
April Meeting of the Metropolitan New York Section		1092
April Meeting of the Southwestern Section		1093
April Meeting of the Texas Section		1093
May Meeting of the Allegheny Mountain Section		1094
May Meeting of the Rocky Mountain Section		1096
May Meeting of the Seaway Section		1097
Calendars of Future Meetings		1098

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss.

Backlog: Main Articles 12 months, Math. Notes 13 months, Research Problems 7 months, Classroom Notes 11 months, Math. Education 10 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to ALEX ROSENBERG, Department of Mathematics, Cornell University, Ithaca, N.Y. 14850; NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D.C. 20036.

HARLEY FLANDERS, *Editor*

ALEX ROSENBERG, *Editor-Elect*

ASSOCIATE EDITORS

JOSHUA BARLAZ	J. G. HARVEY	SEYMOUR SCHUSTER
E. R. BERLEKAMP	ERIC S. LANGFORD	J. ARTHUR SEEBACH, JR.
JANE W. DI PAOLA	P. D. LAX	E. P. STARKE
ROBERT GILMER	ARTHUR MATTUCK	LYNN A. STEEN
RICHARD GUY	M. W. POWNALL	JAMES WENDEL
RAOUL HAILPERN	GIAN-CARLO ROTA	

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are \$12.50. For nonmembers the subscription price is \$18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June-July, August-September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices.

Copyright © The Mathematical Association of America (Incorporated), 1973

PRINTED IN THE UNITED STATES OF AMERICA

THE SPINOR SPANNER

ETHAN D. BOLKER, University of Massachusetts, Boston

1. Introduction. Consider a wrench, which is an object asymmetrical enough so that the result of any proper rotation performed on it is easily recognized. Rotate the wrench through a full 360° turn about an axis. Has it returned to its original state? Physical and geometric intuition both say “yes”, yet the calculus of spinors, which models the quantum mechanical behavior of neutrons, predicts that the answer would be “no” if the wrench were a neutron, or any other Fermion, a particle with half integral spin. More striking still, the predicted answer is “yes” for two full turns (720°) about the same axis. No experiment has yet been performed to verify these predictions, because beam splitters and interferometers for beams of polarized neutrons do not yet exist, but several such experiments have been imagined [1], [2]. There is, however, an easy experiment with an analogous outcome. P. A. M. Dirac invented it to lessen, in lectures, the implausibility of the neutron’s predicted behavior [3]. Consider the wrench again, which Dirac would have called by its English name, a spanner, hence a spinor spanner because of the use to which he put it. Attach it by three cords to the walls of the room. (See the solid lines in Fig. 1.)

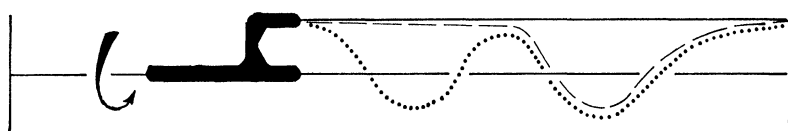


FIG. 1

When we turn the wrench through 360° the cords become tangled (the dashed lines in Fig. 1); no tampering can undo that tangle as long as the wrench is fixed. After two full turns (the dotted line in Fig. 1) the snarl seems worse but is not. Before reading further, find a wrench, perform the experiment, and convince yourself of the striking fact that after two full turns the cords are essentially untangled. The geometry of the spinor spanner is the key to Piet Hein’s topological game Tangloids, described by Martin Gardner in the *Scientific American* [9], and to an ingenious device invented and patented by D. A. Adams which allows a rotating platform to be connected to a stationary base with a flexible cable without using slip rings or rotary joints [8].

I first saw the spinor spanner demonstrated by Norman Ramsey, a physicist, while I was a graduate student. In this paper I shall explain in mathematical terms why the spinor spanner works, and indicate how that explanation can be couched

Ethan Bolker received his Harvard Ph.D. under Andrew Gleason. He has held positions at Princeton, Bryn Mawr College and the University of Massachusetts, Boston. He spent a year’s leave at Berkeley, and a second at Harvard. His main research interest is in combinatorics. He wrote up Professor Lynn Loomis’s lecture notes on *Harmonic Analysis* put out by the M.A.A. in 1965, and he is the author of *Elementary Number Theory, An Algebraic Approach* (W. A. Benjamin, 1970). *Editor.*

in language suitable for mathematics clubs and more general mathematically naive audiences. Someday, I should like to make a movie of the spinor spanner.

We are about to show that the fundamental group G of $SO(3)$, the group of proper rotations of Euclidean 3-space, is of order 2, and to exploit the proof to find a method for untangling the cords. Since Fermions correspond to representations of the double covering group of $SO(3)$ which do not factor through $SO(3)$ itself, the fact that the order of G is 2 really accounts both for the spinor spanner and for the neutron's behavior.

2. Homotopy. Let X be a topological space and x_0 a fixed point in X . A naive audience could think of X as a smooth part of some Euclidean space, say the surface of a sphere, or a solid torus, or an annulus. A *loop* in X is a continuous function $P: [0, 1] \rightarrow X$ for which $P(0) = P(1) = x_0$. If you think of X as a park then a loop may be thought of as the record of an hour's walk in X , starting and ending at x_0 . Be sure to distinguish this precise usage from the more customary meaning of "closed path in a park." The latter is the image of the function P . Two loops P and Q are **homotopic**, written $P \sim Q$, when one can be continuously deformed into the other. Formally, $P \sim Q$ when there is a continuous $f: [0, 1] \times [0, 1] \rightarrow X$ for which $f(0, s) = f(1, s) = x_0$, $f(t, 0) = P(t)$ and $f(t, 1) = Q(t)$. Informally, suppose that you walk your dog in X : you follow P while he follows Q . Then $P \sim Q$ means that when the walk is over the leash joining the two of you can be pulled in without encountering any parklike obstacles, trees or lakes, which you and your dog passed on opposite sides of. This interpretation makes clear the importance of the direction in which you traverse the curve which is the image of P . If P is a sense preserving reparametrization of Q then $P \sim Q$. The loop corresponding to the lazy man's walk is the constant loop 0 defined by $0(t) = x_0$ for all t .

Now let us consider taking two walks in succession. We shall denote " P followed by Q " by " $P \oplus Q$ ". As a function, $P \oplus Q$ is defined by

$$P \oplus Q(t) = \begin{cases} P(2t) & \text{if } 0 < t \leq 1/2 \\ Q(2t - 1) & \text{if } 1/2 \leq t \leq 1. \end{cases}$$

It is intuitively clear and not hard to prove that homotopy is an equivalence relation, that the homotopy class of $P \oplus Q$ depends only on the classes of P and of Q , and that the set of homotopy classes is a group under \oplus . Details can be found in many topology texts (for example, [5] and [7]). The group is not usually abelian, but I have found additive notation less confusing than multiplicative for naive audiences. Observe that 0 is the group identity: $P \oplus 0 \sim P$. Our job now is to find the inverse of P , the solution to $P \oplus ? \sim 0$. The dog walking analogy can lead us to a good guess. If you are lazy while your dog follows P then his leash will be tangled when he returns, unless, by chance, $P \sim 0$. How could you untangle the leash? If the dog is intelligent the answer is easy: ask him to retrace his steps. That is, if we define the loop $-P$ by $(-P)(t) = P(1-t)$ then $P \oplus (-P) \sim 0$.

3. Pasting. The topological spaces we can visualize as smooth parts of 2- or 3-space are too simple to help us analyze the spinor spanner. We need a method for studying homotopy in more complicated ones.

If we take a square piece of paper and paste together a pair of parallel sides, top to top and bottom to bottom, we have made a cylinder. We can study homotopy on the cylinder without actually pasting the square, as long as we remember that points along one edge are identified with corresponding points on the other. The idea of "pasting" can be made precise using quotient topologies, but we have no need for that much sophistication. For naive audiences it is instructive to mention the various spaces which can be obtained by pasting edges of a rectangle. They are the cylinder, the Möbius strip, the torus, the Klein Bottle, and, finally, the projective plane. The identifications which lead to these are symbolically indicated in Figs. 2.1–2.5 respectively, in which some loops are sketched as well. The Klein Bottle and the projective plane cannot actually be constructed in 3-space but we can study them nevertheless. Fig. 2.5 suggests a more symmetrical view of the projective plane Π . Since each pair of opposite points of the square is pasted, the corners assume no special role. We can build Π from a disk Δ by pasting together each pair of antipodal points on the rim: in Fig. 3, these are pairs (A, A') , (B, B') , (C, C') , etc.

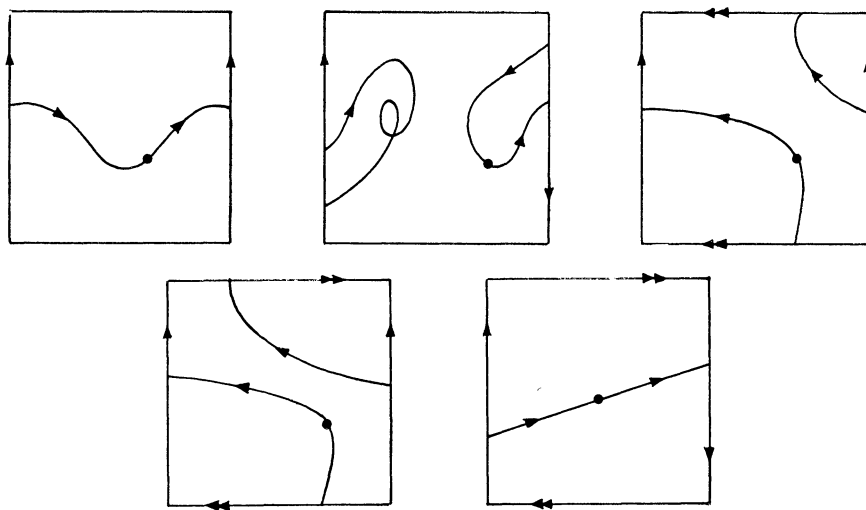


FIG. 2 (2.1–2.5)

Let x_0 be the center and L a directed diameter of Δ . Since the ends of L are identified when we build Π we can consider the loop P in Π which begins at x_0 , follows L to the rim of Δ and then returns to x_0 along the other half of L . We show next that $P \sim 0$; to do so we use a homeomorphic copy or model of Π . Start with the

disk Δ and stretch it to form a closed hemisphere. Now consider the spherical surface Σ of which Δ is a part. If we paste together each pair of antipodal points of Σ , then Π will result. To see this, paste first all the antipodal pairs one member of which lies in the interior of Δ . That yields the hemisphere into which Δ was stretched. The rest of the pasting, of the pairs on the equator, is just what to do to Δ to build Π . In this model for Π the north and south poles n and s of Σ paste together to make x_0 . In Σ there is a unique continuous curve S which starts at n and which becomes P when Σ is pasted to form Π , namely, the appropriate meridian joining n to s . That curve is not a loop in Σ . If P were homotopic to 0 in Π we could lift that homotopy to Σ and so construct a continuous deformation in Σ of S to the constant loop at n during which the endpoints n and s of S remained fixed. Since such a deformation is clearly impossible, $P \sim 0$ in Π . We can see too that $P \oplus P \sim 0$, because $P \oplus P$ is the result in Π of pasting a great circle through n and s in Σ . That great circle easily shrinks to the constant loop at n in Σ . But to untangle cords later, we must now show in another way that $P \oplus P \sim 0$. Consider again our first model for Π , obtained by pasting pairs of opposite points on the rim of Δ .

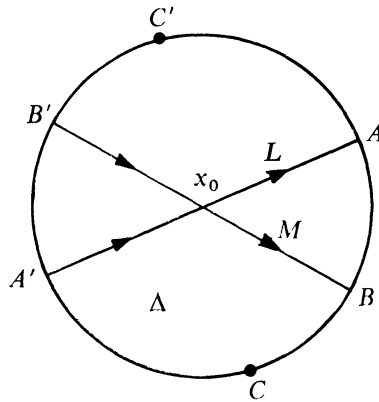


FIG. 3

Let M be another directed diameter of Δ , and let Q follow M in Π as P follows L (see Fig. 3). We can rotate L in Δ until it coincides with M ; this rotation is a continuous deformation in Π of P to Q . If we take for M the diameter L with its direction reversed then Q is $-P$, so $P \oplus P \sim P \oplus (-P) \sim 0$. The projective plane thus surrounds a peculiar kind of hole. If you travel around it twice in the same direction you've not gone around it at all. That is analogous to what happens to the spinor spanner. In each case doubling something makes it vanish. But with the techniques of homotopy and pasting, we can do better than produce an analogy for the spinor spanner. We can predict and explain its behavior.

4. The topology of $SO(3)$. Let Ω be the space of all possible configurations of the wrench. A point $\omega \in \Omega$ is thus the result of a particular proper rotation. Remem-

ber, it is the configuration we are talking about, not the means by which the wrench came to that configuration. It is intuitively clear that Ω is a nice topological space. Our complete turn of the wrench about an axis corresponds to a loop P in Ω which begins and ends at the initial configuration ω_0 . We shall show $P \sim 0$ but $P \oplus P \sim 0$ in Ω and then show how the homotopy which shrinks $P \oplus P$ to 0 tells us how to untangle cords.

We begin by building a model of Ω . Replace the wrench by the surface of a sphere Σ centered at the origin. Then each $\omega \in \Omega$ can be identified with a map from Σ to itself defined by letting $\omega(\sigma) =$ the position of $\sigma \in \Sigma$ when Σ is moved to configuration ω . As a map, ω preserves distances and the orientation of spherical triangles. We next show, in two ways, that every such map has a fixed point. Since ω extends to a proper linear isometry of \mathbb{R}^3 the roots of its cubic characteristic polynomial have product 1 and each is of absolute value 1. Thus those roots are 1, $e^{i\theta}$, $e^{-i\theta}$ for some θ . Since 1 is a root, 1 is an eigenvalue and ω has a fixed point. This argument clearly works in \mathbb{R}^n if and only if n is odd.

Here is a second proof in \mathbb{R}^3 , suitable for audiences who know no linear algebra. Let Σ have circumference 2. For $x, y \in \Sigma$ let $\mu(x, y)$ be the least great circle distance between x and y . The function $f: \Sigma \rightarrow \mathbb{R}$ defined by $f(x) = \mu(x, \omega x)$ is continuous and so assumes its minimum value $\delta \geq 0$ at some $a \in \Sigma$. If $\delta = 0$ then $\omega a = a$ and ω has the fixed point we desire. We shall show next that $\delta > 0$ implies $\delta = 0$. Suppose $\delta > 0$. Since ω is proper it cannot map every point to its antipode. Thus $\delta < 1$, so we can find a hemisphere H containing both a and ωa . In H draw the great circle Γ joining a to ωa ; it has length δ . Now draw two circles C and D centered at a and ωa respectively; make them so small that they lie in H and do not overlap. Let c be the intersection of C and Γ and d the intersection of D and the continuation of Γ . Since $\mu(a, c) = \mu(\omega a, \omega c)$, $\omega c \in D$. But every point on D except d is less than δ units from c , so $\omega c = d$. Now let n and s be the poles for which Γ lies on the equator. Then $\mu(a, n) = \mu(c, n) = 1/2$ so $\mu(\omega a, \omega n) = \mu(\omega c, \omega n) = 1/2$. Therefore $\omega n = n$ or s . But $\omega n = s$ is impossible because ω preserves the orientation of the spherical triangle $a c n$. Thus $\omega n = n$, n is a fixed point, and $\delta = 0$. That is, a must have been fixed to begin with.

Suppose $\omega \neq \omega_0$. Then ω has exactly two fixed points n_ω, s_ω ; which lie at opposite ends of a diameter of Σ , and Σ can be brought to configuration ω by a rotation of r radians about the axis n_ω, s_ω . We wish to consider rotations which are counterclockwise when we look down on n_ω from outer space; this is the familiar right hand rule. In lectures I use an inflatable globe to show a counterclockwise rotation of 102° about the axis joining Bermuda to Perth, Australia, moves Duluth to the Panama Canal. Since a clockwise rotation about an axis is a counterclockwise rotation about the same axis with its north and south poles interchanged, and since rotations through r and $r - 2\pi$ radians about an axis lead to the same ω , we can describe an $\omega \neq \omega_0$ by giving a vector $m(\omega) \neq 0$ with length $\|m(\omega)\| \leq \pi: m(\omega)$

points toward n_ω from the origin, and $\|m(\omega)\| = r$. If we set $m(\omega_0) = 0$, the range of m is the solid ball B of radius π centered at 0. The function which inverts m is one to one except when $\|V\| = \pi$, for rotations through π radians about V and $-V$ lead to the same ω . Thus Ω is modeled by the space X which results when we paste together each pair of antipodal points on the surface of the solid ball B , because $m: \Omega \rightarrow X$ is a homeomorphism; one to one, onto, continuous, and with a continuous inverse. The loop P in X which corresponds to a full turn of the wrench about an axis L starts at the center of B , moves out along L to the surface and returns to the center along the other half of L . It is analogous to the loop with the same name we have just studied in Π . In fact, Π is a subspace of X in a natural way, so that the two loops we have named " P " coincide. Since $P \sim -P$ in Π , $P \sim -P$ in X . For those who like formulas, we give one for that homotopy. Let Δ be the intersection of B with the x, z plane and L the directed diameter which extends to the directed x axis. The homotopy which interests us rotates L in Δ to change P to $-P$. The matrix for a right handed rotation through r radians about the axis in the x, z plane which makes an angle of θ radians with L is

$$f(r, \theta) = \begin{bmatrix} \cos^2 \theta + (\cos r) \sin^2 \theta & -(\sin r) \sin \theta & (1 - \cos r) \sin \theta \cos \theta \\ (\sin r) \sin \theta & \cos r & -(\sin r) \cos \theta \\ (1 - \cos r) \sin \theta \cos \theta & (\sin r) \cos \theta & \sin^2 \theta + (\cos r) \cos^2 \theta \end{bmatrix}.$$

The function f is continuous on $[0, 2\pi] \times [0, \pi]$, $f(\cdot, 0)$ is the loop P , and $f(\cdot, \pi)$ is the loop $-P$, so f is our homotopy in Ω .

To prove $P \sim 0$ in X we cannot merely use the fact that P lives in the subspace Π of X , for although no deformation of P to 0 is possible inside that subspace one might be possible in X . To rule that out we need a new model for Ω analogous to our second model for Π , the one we built by pasting antipodal points on the 2-sphere Σ . Let Φ be the 3-sphere in real 4-space. We can stretch B so that it covers a hemisphere of Φ . Then Ω results when we paste antipodal pairs in Φ , since B results when we paste first those pairs one member of which is interior to B . In this model the north and south poles of Φ paste together to make ω_0 . Now the proof that $P \sim 0$ in Ω proceeds as it did for Π . In technical terms, we have just constructed and then used a simply connected covering space Φ for Ω .

5. Untangling cords. To exploit the fact that $P \sim -P$, and hence that $P \oplus P \sim 0$ in Ω , we must model Ω and loops in it one more way. Consider two concentric spheres; call the inner one the globe (or the wrench, or the neutron) and the outer one the edge of the universe. Suppose the distance between the spheres is 1. Cords, as many of them as we wish to attach, lie initially along radii joining the globe to the edge of the universe. Pack the space between the globe and the edge of the universe with concentric spherical shells Σ_t , where $t \in [0, 1]$ measures the distance of Σ_t from Σ_0 . Each cord is attached to Σ_t where they meet.

Imagine that the shells can slide relative to each other. Let R be any loop in Ω starting and ending at ω_0 ; suppose we manipulate the globe Σ_1 so that at time t it is at $R(t)$. Then the cords cause the intermediate shells to record R : at time t , Σ_t is in configuration $R(t)$ (Σ_t). A homotopy $R \sim Q$ of paths in Ω is a function $f: [0, 1] \times [0, 1] \rightarrow \Omega$ satisfying the conditions listed earlier. If we now manipulate the shells Σ_t so that at time s , shell Σ_t is at position $f(t, s)$ (Σ_t) we shall have deformed the cords, which initially recorded R , to a record of Q . Thus when P is the loop corresponding to a full turn about an axis the homotopy $P \oplus P \sim 0$ really tells us that our cords can be untangled, even if we started with many more than three.

Because $P \sim 0$, no manipulation of the intermediate shells can untangle the cords after one full turn. It is true and slightly subtler that they cannot be untangled at all [6].

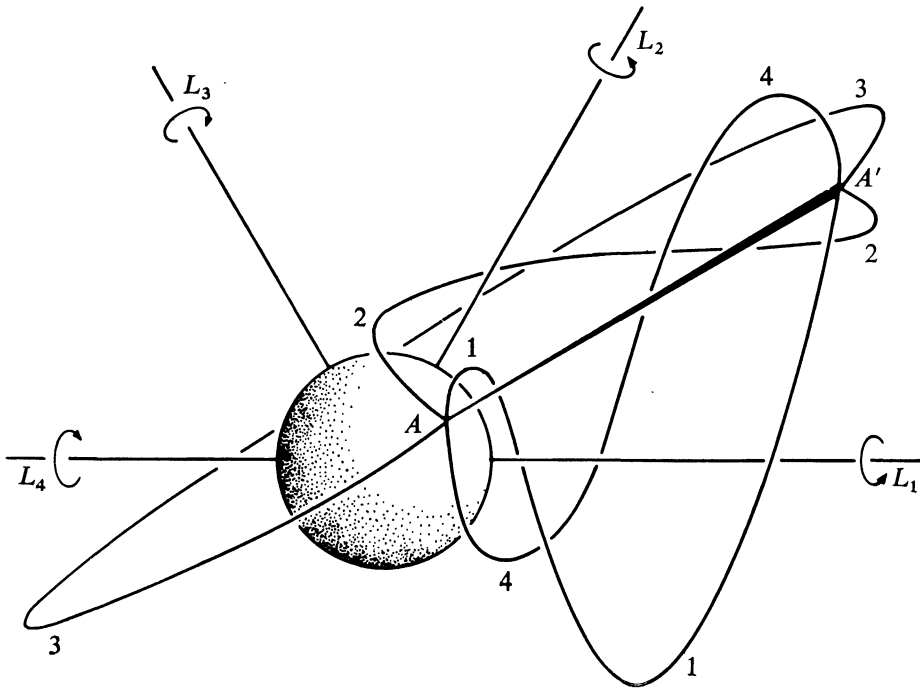


FIG. 4

Let us close by seeing just how the particular homotopy we have studied untangles the cords after two full turns. To convert $P \oplus P$ to 0 we first deform the second summand P to $-P$, or, in other words, deform the result of a full right handed turn about L to the result of a full left handed turn. We do that by rotating L , the axis of the turn, in the subspace Δ of B , so that it reverses its direction. In Fig. 4 we sketch what happens to one of the cords between Σ_1 and $\Sigma_{1/2}$, the one which lies initially

along axis AA' . When the globe executes a full turn about L_i the cord assumes position i . As i varies from 1 to 4, L_i rotates counterclockwise through π radians in the plane of the paper. That operation simultaneously loops the pictured cord and all others on the right over and behind the wrench and those on the left under and in front. That is easier to do than to describe: try it. It really untangles cords. With a little practice it makes a good lecture demonstration or conversation piece, a magic trick which is not magic, but which reflects a fundamental yet little known property of the space in which we live. The analogy between the spinor spanner and the neutron suggests that the state of the latter depends not only on its position and momentum but on which of two topologically distinct ways it is tied to its surroundings. A full turn about an axis leaves its position and momentum unchanged but reverses *its topological relation to the rest of the universe*.

Acknowledgments. Some of the ideas in this paper I explored in conversation with M. Artin, F. C. Cunningham, Jr., M. Gaffney, A. M. Gleason, N. Stein, and N. Ramsey. I am indebted to the Bryn Mawr College Chapter of Sigma Xi and to West Chester State College, West Chester, Pa., where I spoke on the spinor spanner. Final thanks go to Ms. Jessica Bolker, who built the large model wrench I turn while I talk.

References

1. Y. Aharonov and L. Susskind, Observability of the sign change of spinors under 2π rotations, *Phys. Rev.*, 158 (1967) 1237–8.
2. H. J. Bernstein, Can 360° rotations be detected? *Scientific Research* (McGraw-Hill's News Magazine of Science), Vol 4, No. 17 (August 18, 1969) 32–33.
3. P. A. M. Dirac, in conversation on October 30, 1972, remarked that his initial demonstration model was a pair of scissors, to which it is easy to attach the cords.
4. E. Fadell, Homotopy groups of configuration spaces and the string problem of Dirac, *Duke Math. J.*, 29 (1962) 231–242.
5. W. S. Massey, *Algebraic Topology: An Introduction*, Harcourt Brace and World, New York, 1967, chapters 2 and 5.
6. M. H. A. Newman, On the string problem of Dirac, *J. London Math. Soc.*, 17 (1942) 173–177.
7. A. H. Wallace, *An Introduction to Algebraic Topology*, International Series in Pure and Applied Mathematics, Pergamon Press, New York, 1957, Chapter iv.
8. D. A. Adams, Apparatus for Providing Energy Communication Between a Moving and a Stationary Terminal, U.S. Patent 3,586,413, June 22, 1971. Copies available from the U.S. Patent Office or from Mr. Adams at 7434 E. Montecito Drive, Tucson, Ariz. 85710.
9. M. Gardner, *New Mathematical Diversions from Scientific American*, Simon and Schuster, New York, 1966, Chapter 2.

LINEAR COMBINATIONS OF SETS OF CONSECUTIVE INTEGERS

D. A. KLARNER, Stanford University, and R. RADO, University of Reading, England

Dedicated to Paul Erdős on his sixtieth birthday.

Let $k - 1, m_1, \dots, m_k$ denote positive integers such that m_1, \dots, m_k have greatest common divisor 1, and let t denote an integer. A well-known result in the elementary theory of numbers is that the equation

$$(1) \quad m_1 x_1 + \dots + m_k x_k = t$$

has infinitely many solutions in integers x_1, \dots, x_k . Furthermore, there exists an integer $\sigma(\bar{m})$ which depends on $\bar{m} = (m_1, \dots, m_k)$ such that (1) has a solution in non-negative integers x_1, \dots, x_k for all $t \geq \sigma(\bar{m})$, but no solution of this kind exists when $t = \sigma(\bar{m}) - 1$. In this note we prove a refinement of this result by showing that a set of consecutive integers can be obtained by allowing the x_i in (1) to range over suitable sets of consecutive integers. For example, every number t with $6 \leq t \leq 11$ can be expressed in the form $3x + 4y$ with $0 \leq x \leq 3, 0 \leq y \leq 2$. Later on we express facts like this by writing

$$(2) \quad [6, 11] \subseteq 3[0, 3] + 4[0, 2].$$

The following notation is used: I, N , and P denote the set of all integers, the set of all non-negative integers, and the set of all positive integers respectively. Also, for any pair of elements $i, j \in I$, define $[i, j] = \{x: x \in I, i \leq x \leq j\}$; furthermore, given sets $I_1, \dots, I_k \subseteq I$ together with elements $m_1, \dots, m_k \in I$, define

$$(3) \quad m_1 I_1 + \dots + m_k I_k = \{m_1 x_1 + \dots + m_k x_k: x_i \in I_i \ (i = 1, \dots, k)\}.$$

For each $k \in P$ and $J \subseteq I$, let J^k denote the set of all k -dimensional vectors over J ; next, for elements $\bar{x}, \bar{y} \in I^k$ with $\bar{x} = (x_1, \dots, x_k), \bar{y} = (y_1, \dots, y_k)$ define the usual dot product $\bar{x} \cdot \bar{y} = x_1 y_1 + \dots + x_k y_k$; finally, define $\bar{x} < \bar{y}$ whenever $x_i < y_i$ for $i = 1, \dots, k$, and define $\bar{x} \leq \bar{y}$ whenever $x_i \leq y_i$ for $i = 1, \dots, k$.

Our main result may be succinctly stated in this notation as follows.

THEOREM 1. Suppose $k - 1, m_1, \dots, m_k \in P$ and m_1, \dots, m_k have greatest common divisor 1; let $\bar{m} = (m_1, \dots, m_k)$ and $m = \max \{m_1, \dots, m_k\}$; suppose $\bar{u}, \bar{v} \in I^k$ satisfy

$$(4) \quad \bar{v} - \bar{u} \geq (m - 1, \dots, m - 1),$$

$$(5) \quad \bar{m} \cdot (\bar{v} - \bar{u}) > 2(m - 1)(m_1 + \dots + m_k).$$

Then

$$(6) \quad [\bar{m} \cdot \bar{u} + \sigma(\bar{m}), \bar{m} \cdot \bar{v} - \sigma(\bar{m})] \subseteq m_1[u_1, v_1] + \dots + m_k[u_k, v_k],$$

where $\bar{u} = (u_1, \dots, u_k), \bar{v} = (v_1, \dots, v_k)$, and $\sigma(\bar{m})$ is the function defined after (1).

Before proving Theorem 1, we shall state and prove a result dealing with the 2-dimensional situation which is sharper than the result provided by taking $k = 2$ in Theorem 1. Furthermore, the proof of Theorem 2 gives some insight for the proof of Theorem 1.

THEOREM 2: *Suppose $m_1, m_2 \in P$ such that m_1 and m_2 are relatively prime; also, suppose $u_1, u_2, v_1, v_2 \in I$ such that $v_1 - u_1 \geq m_2 - 1$, $v_2 - u_2 \geq m_1 - 1$. Then*

$$(7) \quad [m_1 u_1 + m_2 u_2 + (m_1 - 1)(m_2 - 1), m_1 v_1 + m_2 v_2 - (m_1 - 1)(m_2 - 1)] \\ \subseteq m_1 [u_1, v_1] + m_2 [u_2, v_2].$$

Proof: It is well known that $\sigma(m_1, m_2) = (m_1 - 1)(m_2 - 1)$, where $\sigma(m_1, m_2) - 1$ denotes the largest integer not expressible in the form $m_1 x + m_2 y$ with $x, y \in N$. Let $\bar{m} = (m_1, m_2)$, $\bar{u} = (u_1, u_2)$, and $\bar{v} = (v_1, v_2)$, then it follows from the definition of $\sigma(\bar{m})$ that

$$(8) \quad \bar{m} \cdot \bar{u} + \sigma(\bar{m}) + N \subseteq m_1(u_1 + N) + m_2(u_2 + N),$$

$$(9) \quad \bar{m} \cdot \bar{v} - \sigma(\bar{m}) - N \subseteq m_1(v_1 - N) + m_2(v_2 - N).$$

Hence, the intersection of the sets on the left in (8) and (9) is contained in the intersection of the sets on the right in (8) and (9). That is,

$$(10) \quad [\bar{m} \cdot \bar{u} + \sigma(\bar{m}), \bar{m} \cdot \bar{v} - \sigma(\bar{m})] \\ \subseteq (m_1(u_1 + N) + m_2(u_2 + N)) \cap (m_1(v_1 - N) + m_2(v_2 - N)).$$

Now we prove a remarkable identity which gives a valid instance of intersection distributing over addition.

$$(11) \quad (m_1(u_1 + N) + m_2(u_2 + N)) \cap (m_1(v_1 - N) + m_2(v_2 - N)) \\ = m_1((u_1 + N) \cap (v_1 - N)) + m_2((u_2 + N) \cap (v_2 - N)).$$

Of course, the set on the right in (11) is just

$$(12) \quad m_1 [u_1, v_1] + m_2 [u_2, v_2],$$

so (10), (11), and (12) combine to imply (7). It remains to prove (11).

Consider the set of points $I \times I$ in the Cartesian plane. The subsets $(u_1 + N) \times (u_2 + N)$ and $(v_1 - N) \times (v_2 - N)$ of $I \times I$ lie in upper and lower quadrants of the plane whose intersection contains the set $[u_1, v_1] \times [u_2, v_2]$. This situation is illustrated in Figure 1. We want to study the linear form $m_1 x + m_2 y$ evaluated over all points $(x, y) \in I \times I$; in particular, we are interested in points which have equal evaluations. Given an element $h \in I$, the set L_h of all points $(x, y) \in I \times I$ such that $m_1 x + m_2 y = h$ is situated on a unique line having slope $-m_1/m_2$. Also, it is easy to see that if $(x', y') \in (I \times I) \cap L_h$, then $L_h = \{(x' + jm_2, y' - jm_1) : j \in I\}$.

To prove (11), note that the set on the right is contained in the set on the left;

suppose the reverse is not true. From this assumption we shall deduce a contradiction. Under this assumption it follows that there exists an $h \in I$ such that L_h has points in common with both

$$U = ((u_1 + N) \times (u_2 + N)) \setminus ([u_1, v_1] \times [u_2, v_2])$$

and

$$V = ((v_1 - N) \times (v_2 - N)) \setminus ([u_1, v_1] \times [u_2, v_2]),$$

but L_h has no point in common with $B = [u_1, v_1] \times [u_2, v_2]$.

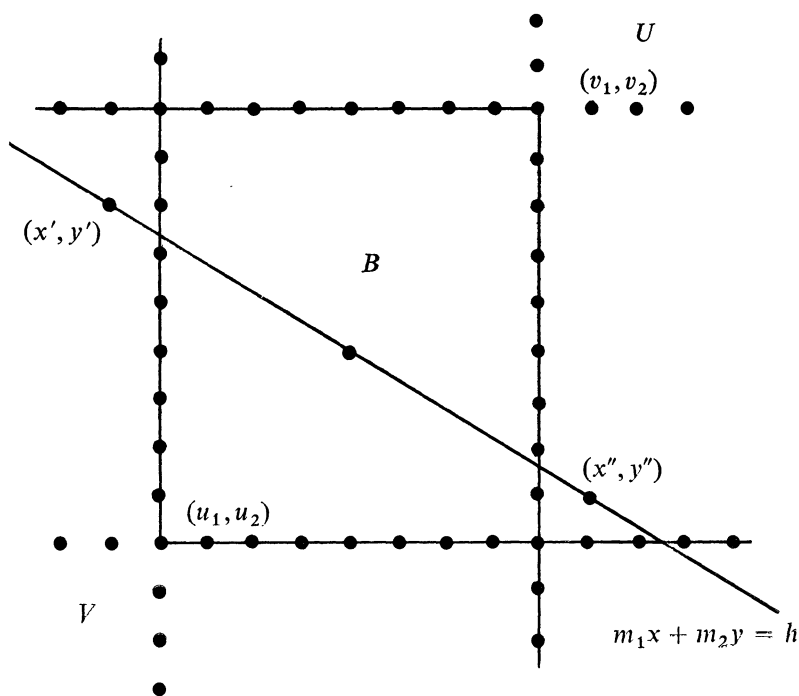


FIG. 1. The set of points $(u_1 + N) \times (u_2 + N)$ lies in the quadrant above and to the right of the point (u_1, u_2) , the set of points $(v_1 - N) \times (v_2 - N)$ lies in the quadrant below and to the left of the point (v_1, v_2) , and the set of points $[u_1, v_1] \times [u_2, v_2]$ lies in the box.

Suppose $(x', y') \in L_h \cap U$ and $(x'', y'') \in L_h \cap V$; since $(x', y') \notin B$, either $x' < u_1$ or $y' > v_2$. If $x' < u_1$, then $x'' > v_1$ because $(x', y'), (x'', y'') \in L_h$ and $(x'', y'') \notin B$. In this case we suppose (x', y') has been selected from $L_h \cap U$ so that x' is maximal, and (x'', y'') has been selected from $L_h \cap V$ so that x'' is minimal. Since $(x', y'), (x'', y'') \in L_h$, and $L_h \cap B = \emptyset$, we must have $x'' - x' = m_2$. But, $x' < u_1$ and $x'' > v_1$ implies $x' + 1 \leq u_1$ and $x'' - 1 \geq v_1$; hence, $m_2 - 2 = x'' - x' - 2 \geq v_1 - u_1$, contradicting the hypothesis $v_1 - u_1 \geq m_2 - 1$. In the case $y' > v_2$, it follows that $y'' < u_2$. This time the points (x', y') and (x'', y'') are selected so that y' is minimal and y'' is maximal.

The argument goes just as before; we must have $y' - y'' = m_1$ which leads to the contradiction $v_2 - u_2 \leq m_1 - 2$. This completes the proof of Theorem 2.

Now we prove Theorem 1. To do this, we prove an identity having the form of (11), but subject to the conditions (4) and (5).

LEMMA. If k -dimensional vectors \bar{m}, \bar{u} , and \bar{v} satisfy the hypothesis of Theorem 1, then

$$(13) \quad \sum_{i=1}^k m_i(u_i + N) \cap \sum_{i=1}^k m_i(v_i - N) = \sum_{i=1}^k m_i((u_i + N) \cap (v_i - N)).$$

Theorem 1 is an immediate consequence of the Lemma; its application is the justification of the penultimate equality in the following string of formulas.

$$(14) \quad \begin{aligned} [\bar{m} \cdot \bar{u} + \sigma(\bar{m}), \bar{m} \cdot \bar{v} - \sigma(\bar{m})] &= (\bar{m} \cdot \bar{u} + \sigma(\bar{m}) + N) \cap (\bar{m} \cdot \bar{v} - \sigma(\bar{m}) - N) \\ &\subseteq \sum_{i=1}^k m_i(u_i + N) \cap \sum_{i=1}^k m_i(v_i - N) \\ &= \sum_{i=1}^k m_i((u_i + N) \cap (v_i - N)) = \sum_{i=1}^k m_i[u_i, v_i]. \end{aligned}$$

To prove Theorem 1 completely, it remains to prove the Lemma. For each $i \in I$, let $L_i = \{\bar{x}: \bar{x} \in I^k; \bar{m} \cdot \bar{x} = i\}$, and suppose the Lemma is false. Then there exists $h \in I$ such that $L_h \cap U, L_h \cap V \neq \emptyset$, but $L_h \cap B = \emptyset$ where

$$\begin{aligned} U &= \{\bar{x}: \bar{x} \in I^k, \bar{x} \geq \bar{u}\} \setminus B \\ V &= \{\bar{x}: \bar{x} \in I^k, \bar{x} \leq \bar{v}\} \setminus B \\ B &= [u_1, v_1] \times \cdots \times [u_k, v_k]. \end{aligned}$$

Suppose $\bar{x}' \in U$ is selected so that

$$(15) \quad \sum_{i=1}^k \max\{v_i, x'_i\}$$

is minimal, where $\bar{x}' = (x'_1, \dots, x'_k)$. Since $\bar{x}' \notin B$, there exists $r \in [1, k]$ such that $x'_r > v_r$. Furthermore, there exists $s \in [1, k]$ such that $x'_s \leq v_s$ since otherwise $\bar{x}' > \bar{v}$, which implies $h = \bar{m} \cdot \bar{x}' > \bar{m} \cdot \bar{v}$ for all $\bar{x} < \bar{v}$, contradicting the assumption $L_h \cap V \neq \emptyset$. Of course, $r \neq s$, so we have

$$(16) \quad h = \sum_{\substack{i=1 \\ i \neq r, s}}^k m_i x'_i + m_r(x'_r - m_s) + m_s(x'_s + m_r);$$

$$(17) \quad \begin{aligned} x'_r - m_s - u_r &\geq (v_r + 1) - m_s - u_r = (v_r - u_r) - m_s + 1 \\ &\geq (v_r - u_r) - m + 1 \geq 0. \end{aligned}$$

Hence, by the minimality assumption made in (15),

$$(18) \quad \max \{v_r, x'_r - m_s\} + \max \{v_s, x'_s + m_r\} \geq \max \{v_r, x'_r\} + \max \{v_s, x'_s\}.$$

Hence,

$$(19) \quad \begin{aligned} \max \{v_s, x'_s + m_r\} &> \max \{v_s, x'_s\} = v_s; \\ x'_s + m_r &> v_s; \quad x'_s > v_s - m_r \geq v_s - m. \end{aligned}$$

This implies

$$(20) \quad \bar{x}' > \bar{v} - (m, \dots, m).$$

Suppose $x'' \in V$ is selected so that

$$(21) \quad \sum_{i=1}^k \min \{u_i, x''_i\}$$

is maximal where $\bar{x}'' = (x''_1, \dots, x''_k)$. Now an argument running parallel to (15)–(21) can be given to show that

$$(22) \quad \bar{x}'' < \bar{u} + (m, \dots, m).$$

Together (20) and (22) imply

$$(23) \quad \begin{aligned} 0 &= \bar{m} \cdot \bar{x}' - \bar{m} \cdot \bar{x}'' \geq \sum_{i=1}^k m_i ((v_i - m + 1) - (u_i + m - 1)) \\ &= \bar{m} \cdot (\bar{v} - \bar{u}) - 2(m - 1) \sum_{i=1}^k m_i. \end{aligned}$$

But (5) implies

$$(24) \quad \bar{m} \cdot (\bar{v} - \bar{u}) - 2(m - 1) \sum_{i=1}^k m_i > 0,$$

so (23) provides the required contradiction, and we conclude that the Lemma is true.

The results proved in this paper arose in connection with our investigation [1] of the smallest set $\langle \bar{m} \cdot \bar{x} : 1 \rangle \subseteq P$ containing 1 which is closed under the operation $\bar{m} \cdot \bar{x}$, where $\bar{m} = (m_1, \dots, m_k)$ is a given k -tuple of relatively prime positive integers.

This research was supported by the Office of Naval Research under contract number N-00014-67-A-0112-0057 NR 044-402, and by the National Science Foundation under grant number GJ-002. Reproduction in whole or in part is permitted for any purpose of the United States Government.

At the time this paper was written, R. Rado was Visiting Professor at the Faculty of Mathematics, Department of Combinatorics and Optimization, University of Waterloo, Ontario, Canada.

Reference

1. D. A. Klarner and R. Rado, Arithmetic Properties of Certain Recursively Defined Sets, to appear.

THE EQUATION $x'(t) = ax(t) + bx(t - \tau)$ WITH "SMALL" DELAY

R. D. DRIVER, University of Rhode Island,
D. W. SASSER, Sandia Laboratories, Albuquerque,
M. L. SLATER, Texas Christian University

One of the simplest examples of a delay differential equation is the linear scalar equation

$$(1) \quad x'(t) = ax(t) + bx(t - \tau),$$

where $a, b \neq 0$, and $\tau > 0$ are real constants. As is well known and easily proved, for every given function $\phi \in C([-\tau, 0], R)$, there exists a unique function $x \in C([-\tau, \infty), R)$ which satisfies the initial condition

$$(2) \quad x(t) = \phi(t) \text{ for } t \in [-\tau, 0]$$

and which satisfies Eq. (1) for $t > 0$. We shall call this function x the solution of Eq. (1) with initial condition (2) or, more briefly, the solution of Eqs. (1) and (2).

Equation (1) occurs in a number of applications. For example, a certain model for population growth gives rise to the nonlinear equation

$$x'(t) = -cx(t - 1)[1 + x(t)].$$

Here the population is proportional to $1 + x(t)$. The same nonlinear equation has even arisen in the study of the distribution of prime numbers. The stability of the trivial solution of this nonlinear equation depends upon the stability of the trivial solution of its linear approximation

$$x'(t) = -cx(t - 1),$$

a special case of (1). See, for example, Wright [11].

Another equation which has been proposed as a model for population growth, and also for gonorrhea epidemiology, is

$$x'(t) = g(x(t)) - g(x(t - L)).$$

The linear approximation to this equation is Eq. (1) with $a + b = 0$. See Cooke and Yorke [2].

As a third application, consider the problem of mixing of salt brines encountered in any elementary differential equations text. The usual example assumes that an inflowing salt solution is instantaneously perfectly mixed with the brine in the tank. The mixture simultaneously flows out at the bottom of the tank at the same rate as the inflow. The resulting differential equation for the amount of salt in the tank is $y'(t) = k - cy(t)$. But if one eliminates the assumption of instantaneous perfect mixing, he is naturally led [4] to the equation

$$y'(t) = k - cy(t - \tau).$$

It now suffices to introduce $x(t) \equiv y(t) - k/c$ and the equation becomes $x'(t) = -cx(t - \tau)$, a special case of (1) again.

The problem represented by Eqs. (1) and (2) has received much study. The known results for this problem include series expansions of the solutions, due to Schmidt [9] and others, (see Bellman and Cooke [1] or El'sgol'ts [5]); a detailed study of the asymptotic behavior of solutions by Myškis [6]; and an asymptotic characterization of the solutions in case the delay is "small", due to Rjabov [8] (see also [3]). Actually equation (1) is merely a simple prototype of the various equations considered by these authors.

The present note, restricted to (1), shows how some of the known results can be simply obtained using elementary calculus. When the delay τ is small, we shall prove that certain similarities exist between the solutions of Eq. (1) and those of an equation without delay.

THEOREM. *Let*

$$(3) \quad -\frac{1}{e} < b\tau e^{-a\tau} < e.$$

Then, in the real interval $(a - 1/\tau, \infty)$, the characteristic equation

$$(4) \quad \lambda = a + be^{-\lambda\tau}$$

has a unique solution λ . Moreover $\lambda < a + 1/\tau$. And, if λ is this particular solution of (4), and if x is the solution of Eqs. (1) and (2), then

$$(5) \quad \lim_{t \rightarrow \infty} [x(t)e^{-\lambda t}] = \frac{1}{1 + b\tau e^{-\lambda\tau}} [\phi(0) + be^{-\lambda\tau} \int_{-\tau}^0 e^{-\lambda s} \phi(s) ds],$$

the limit being approached exponentially.

REMARK. For an equation with several delays,

$$x'(t) = ax(t) + \sum_{j=1}^m b_j x(t - \tau_j),$$

where $0 \leq \tau_j \leq \tau$ for $j = 1, \dots, m$, a similar result holds under the hypothesis

$$\tau \sum_{j=1}^m |b_j| e^{(-a+1/\tau)\tau_j} < 1.$$

For Eq. (1), however, this condition is much stricter than (3). The case of infinitely many distributed delays has been treated elsewhere [4].

Proof of the Theorem. To analyze the characteristic equation (4), let us consider the function D , defined by $D(p) = p - a - be^{-p\tau}$. It follows from the first inequality of (3) that

$$D\left(a - \frac{1}{\tau}\right) = -\frac{1}{\tau} - be^{-a\tau+1} < 0.$$

Again using the first inequality of (3), we find that for all $p \geq a - 1/\tau$

$$D'(p) = 1 + b\tau e^{-p\tau} > 1 - e^{a\tau-1}e^{-p\tau} \geq 0.$$

Since $\lim_{p \rightarrow \infty} D(p) = \infty$, it follows that there is a unique $\lambda > a - 1/\tau$ such that $D(\lambda) = 0$.

Invoking the second inequality of (3), we find that

$$D\left(a + \frac{1}{\tau}\right) = \frac{1}{\tau} - be^{-a\tau-1} > 0.$$

Thus it follows that

$$\lambda \in \left(a - \frac{1}{\tau}, a + \frac{1}{\tau}\right).$$

This in turn enables us to estimate

$$(6) \quad |b\tau e^{-\lambda\tau}| = |\lambda - a|\tau < 1.$$

Now define $y(t) = x(t)e^{-\lambda t}$ and find the equations, equivalent to (1) and (2), for y :

$$(7) \quad y'(t) = -be^{-\lambda\tau}[y(t) - y(t - \tau)] \text{ for } t > 0,$$

with the initial condition

$$(8) \quad y(t) = \phi(t)e^{-\lambda t} \text{ for } -\tau \leq t \leq 0.$$

These equations, in turn, are equivalent to

$$(9) \quad y(t) = -be^{-\lambda\tau} \int_{t-\tau}^t y(s) ds + C \text{ for } t \geq 0$$

with (8), where

$$(10) \quad C = \phi(0) + be^{-\lambda\tau} \int_{-\tau}^0 \phi(s)e^{-\lambda s} ds.$$

Since $1 + b\tau e^{-\lambda\tau} > 0$, we can define

$$z(t) = y(t) - \frac{C}{1 + b\tau e^{-\lambda\tau}}$$

and get another equivalent problem:

$$(11) \quad z(t) = -be^{-\lambda\tau} \int_{t-\tau}^t z(s) ds \text{ for } t \geq 0$$

with

$$(12) \quad z(t) = \phi(t)e^{-\lambda t} - \frac{C}{1 + b\tau e^{-\lambda\tau}} \text{ for } -\tau \leq t \leq 0.$$

Let M be an upper bound for $|z(t)|$ on $[-\tau, 0]$. Then we shall show that $|z(t)| \leq M$ for all $t \geq -\tau$. Given any $\varepsilon > 0$, suppose (for contradiction) that $|z(t)| < M + \varepsilon$ for $-\tau \leq t < t_1$ and $|z(t_1)| = M + \varepsilon$. Then, using (6) for the first time, we find

$$M + \varepsilon = |z(t_1)| \leq |be^{-\lambda\tau}| \int_{t_1-\tau}^{t_1} |z(s)| ds \leq |be^{-\lambda\tau}| \tau(M + \varepsilon) < M + \varepsilon,$$

which is nonsense. Thus $|z(t)| < M + \varepsilon$ for all $t \geq -\tau$, and, since ε was arbitrary, $|z(t)| \leq M$ for all $t \geq -\tau$.

It now follows, by an easy induction, that $|z(t)| \leq |b\tau e^{-\lambda\tau}|^n M$ for all $t \geq n\tau - \tau$ ($n = 0, 1, 2, \dots$), and hence $z(t) \rightarrow 0$ exponentially as $t \rightarrow \infty$. This is equivalent to (5).

As has already been indicated, the result of this theorem is not really new. It can essentially be found, using Laplace-transform and residue-theory methods, in Chapter 4 of [1] or Chapter 2 of [5], for example. A similar qualitative result is obtained for much more general equations by Myškis [6] (Chapters 3 and 4) using comparison techniques, and by Rjabov [8] and Uvarov [10] (see also [3]). However, in the latter works, the *value* of the limit of $x(t)e^{-\lambda t}$ (the right hand side of (5)) is not obtained.

The thing which is apparently new here is the simple proof of equation (5), using nothing but elementary calculus.

The theorem provides the asymptotic behavior of $x(t)$, except in the case when

$$(13) \quad \phi(0) + be^{-\lambda\tau} \int_{-\tau}^0 e^{-\lambda s} \phi(s) ds = 0.$$

And a randomly-chosen function $\phi \in C([-\tau, 0], \mathbb{R})$ will rarely satisfy (13). More precisely, one can show that equation (13) is satisfied only when ϕ belongs to a certain nowhere dense subset of $C([-\tau, 0], \mathbb{R})$ with the sup norm.

M. J. Norris [7] has independently considered an equation like (1) but with infinitely many (distributed) delays. Assuming negative coefficients and a sufficiently small maximum delay, he has determined a *two-term* asymptotic representation; and, for the special case of (1), his proof is also quite elementary.

We conclude by giving, as corollaries, some easy consequences of the theorem proved here.

COROLLARY 1. *If $-1/e < b\tau e^{-a\tau} < e$, then equation (1) has no oscillatory solution except possibly in the unlikely case that (13) holds.*

COROLLARY 2. *Let $-1/e < b\tau e^{-a\tau} < e$. Then*

- (i) $\lambda < 0$ whenever $a + b < 0$ and $a\tau < 1$,
- (ii) $\lambda = 0$ whenever $a + b = 0$ and $a\tau < 1$,
- (iii) $\lambda > 0$ whenever either $a + b > 0$ or $a\tau > 1$.

The trivial solution of (1) is (uniformly) asymptotically stable in case (i), (uniformly) stable in case (ii), and unstable in case (iii).

Proof. Whenever $a\tau < 1$, it follows that $b\tau > -e^{a\tau-1} > -1$. Referring to the proof of the theorem we find:

(i) If $a + b < 0$ and $a\tau < 1$, then $D(0) > 0$ and $0 > a - 1/\tau$. Thus $\lambda < 0$.

(ii) If $a + b = 0$ and $a\tau < 1$, then $D(0) = 0$ and $0 > a - 1/\tau$. Thus $\lambda = 0$.

(iii) If $a + b > 0$, then $D(0) = -a - b < 0$. Thus $\lambda > 0$.

And if $a\tau > 1$, then $a - 1/\tau > 0$. Thus $\lambda > 0$ again.

The three stability assertions now follow easily from equation (5).

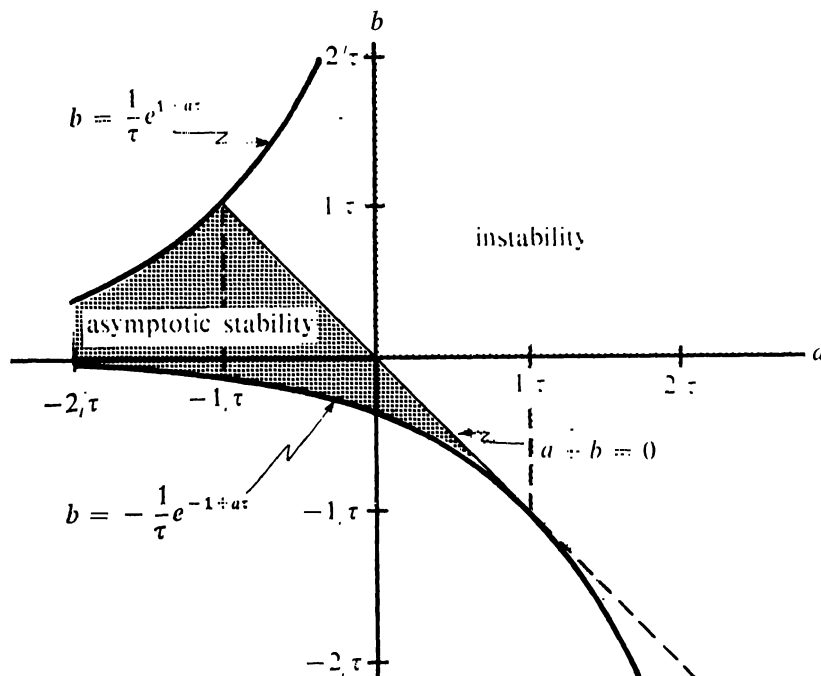
COROLLARY 3. Let $-1/e < b\tau e^{-a\tau} < e$ and $a\tau < 1$. Then the asymptotic behavior of solutions of (1) (assuming (13) is not satisfied) is qualitatively the same as that of the ordinary differential equation obtained by either ignoring the delay,

$$y'(t) = ay(t) + by(t),$$

or by approximating $x(t - \tau)$ with the first two terms of a Taylor's series

$$y'(t) = ay(t) + by(t) - b\tau y'(t).$$

The various regions of the (a, b) -plane, mentioned above, are indicated in the following figure. A complete stability diagram can be found in El'sgol'ts [5], p. 56 (where our a and b are replaced by $-a$ and $-b$).



COROLLARY 4. *Let $a + b = 0$ and $-1 < a\tau < 1$. Then the solution x of (1) and (2) approaches a limit as $t \rightarrow \infty$:*

$$x(t) \rightarrow \frac{1}{1 + b\tau} \left[\phi(0) + b \int_{-\tau}^0 \phi(s) ds \right].$$

Presented to the American Mathematical Society, January 17, 1972 at Las Vegas. This work was partially supported by the United States Atomic Energy Commission.

References

1. R. Bellman and K. L. Cooke, *Differential-Difference Equations*, Academic Press, New York, 1963. MR 26 # 5259.
2. K. L. Cooke and J. A. Yorke, *Equations modelling population growth, economic growth, and gonorrhea epidemiology*, *Ordinary Differential Equations*, Academic Press, New York, 1972, 35–53.
3. R. D. Driver, On Ryabov's asymptotic characterization of the solutions of quasi-linear differential equations with small delays, *SIAM Rev.*, 10 (1968) 329–341. MR 38 # 2410.
4. ———, Some harmless delays, *Delay and Functional Differential Equations and their Applications*, Academic Press, New York, 1972, 103–119.
5. L. E. El'sgol'ts, *Introduction to the Theory of Differential Equations with Deviating Arguments*, Holden-Day, San Francisco, 1966. MR 33 # 381.
6. A. D. Myškis, *Linear Differential Equations with Retarded Argument* (Russian), GITTL, Moscow, 1951. MR 14–52.
7. M. J. Norris, unpublished notes.
8. Ju. A. Rjabov, Certain asymptotic properties of linear systems with small time lag (Russian), *Trudy Sem. Teor. Differencial. Uravnenii s Otklon. Argumentom Univ. Družby Narodov Patrisa Lumumby* 3 (1965) 153–164. MR 35 # 1895.
9. E. Schmidt, Über eine Klasse linearer funktionaler Differentialgleichungen, *Math. Ann.*, 70 (1911) 499–524.
10. V. B. Uvarov, Asymptotic properties of the energy distribution of neutrons slowed down in an infinite medium (Russian), *Ž. Vyčisl. Mat. i. Mat. Fiz.*, 7 (1967) 836–851.
11. E. M. Wright, A non-linear difference-differential equation, *J. Reine Angew. Math.*, 194 (1955) 66–87. MR 17–272.

THE QUATERNION CALCULUS

C. A. DEAVOURS, The Cooper Union of New York
(Current address: Newark State College, Union, N.J.)

1. Introduction. Most students, upon completing a first course in complex analysis, have glimpsed the immense power and elegance of the subject, particularly in treating two dimensional physical problems. The question then arises as to whether an analogous calculus exists for three dimensions. Lack of an appropriate hyper-complex number system seems to prevent any attempt along this line from going

Cipher Deavours received his University of Virginia Sc. D. under Gordon Latta. Since then he has been at The Cooper Union. His main research interest is ordinary differential equations. *Editor.*

very far. Nevertheless, there exists an extensively developed four dimensional calculus, little known in this country, which was developed by R. Fueter [1] in the decade following 1935. A good bibliography to papers on the subject is found in [2]. Rose's work on quaternion velocity potentials for axisymmetric fluid flow appears to be the only paper on the subject to appear in English.

Fueter defines both right and left regular functions of a quaternion variable and develops the associated theory by producing analogues of both Cauchy Theorems, Liouville's Theorem, and Laurent series developments. In quaternion [4] Abelian functions having four periods are constructed and their properties studied.

Some of the essential aspects of Fueter's calculus will be discussed in this paper, using a somewhat different approach. The author has found that selected topics from this subject provide excellent optional topics for courses in complex variables, especially for the more enquiring students. Once acquainted with quaternions students often guess and prove theorems analogous to those which they have recently learned in the course. Science and engineering students gain greatly from such exposure as the use of quaternions provides them with a "concrete" example of an algebra more complicated than that of ordinary complex numbers. (Students never seem to view matrices in this manner.)

The compact quaternion form of Maxwell's equations which has been discovered repeatedly by undergraduates over the years is included along with several other topics of classroom interest.

2. Quaternions. Quaternions were invented in 1843 by the Irish mathematician William Rowen Hamilton after a lengthy struggle to extend the theory of complex numbers to three dimensions. An account of Hamilton's ultimate rejection of the commutative law of multiplication and the ensuing quaternion wars which raged afterwards is to be found in [5] and [6].

The algebra of quaternions has the distinction of being one of the three associative division algebras possible. Linear combinations are formed of the four units 1, **i**, **j**, **k** using coefficients taken from the real number field. The quaternion thus formed, $w + xi + yj + zk$ will be denoted **q** or $w + \mathbf{r}$, where **r** is the usual radius vector of three dimensions. The w component of **q** is called the **scalar part** of the quaternion and **r** is termed its **vector part**. Quaternion addition and scalar multiplication are defined in the usual manner as to constitute a linear algebra. The symbol 1 behaves as the ordinary number one in multiplication while the other units satisfy: $i^2 = j^2 = k^2 = -1$, $ij = k = -ji$, $jk = i = -kj$, $ki = j = -ik$. Products of quaternions are formed using the above rules and the distributive law. Thus

$$(a + \mathbf{A})(b + \mathbf{B}) = ab - \mathbf{A} \cdot \mathbf{B} + a\mathbf{B} + b\mathbf{A} + \mathbf{A} \times \mathbf{B},$$

where the dot and cross indicate the usual three dimensional scalar and vector cross products respectively. For any quaternion $\mathbf{q} = w + \mathbf{r}$ there exists a **conjugate quaternion**, $\bar{\mathbf{q}} = w - \mathbf{r}$, satisfying $\mathbf{q}\bar{\mathbf{q}} = \bar{\mathbf{q}}\mathbf{q} = w^2 + x^2 + y^2 + z^2 = |\mathbf{q}|^2$. The non-

negative quantity $|\mathbf{q}|$ is termed the **norm** of \mathbf{q} . The conjugation operation satisfies the equation $\overline{\mathbf{A}\mathbf{B}} = \overline{\mathbf{B}}\overline{\mathbf{A}}$. Quaternion multiplication is not commutative but all other algebraic properties of the real and complex numbers hold.

The skew field of quaternions is isomorphic to a subset of 4 by 4 matrices under the mapping:

$$\mathbf{q} \rightarrow \begin{bmatrix} w & x & y & z \\ -x & w & -z & y \\ -y & z & w & -x \\ -z & -y & x & w \end{bmatrix}$$

or to a set of 2 by 2 complex matrices related to the Pauli spin matrices [8]. The topological properties of the quaternion group are discussed in [7]. We shall consider functions of a quaternion variable \mathbf{q} which will be written $\mathbf{F}(\mathbf{q})$; such functions can be decomposed into a scalar and vector part which we shall write as $\mathbf{F}(\mathbf{q}) = \phi + \psi$. The vector part of \mathbf{F} will be expressed in component form as $\psi = \psi_1 \mathbf{i} + \psi_2 \mathbf{j} + \psi_3 \mathbf{k}$. Generally, the four components of \mathbf{F} will be required to possess continuous partial derivatives up to a certain order, usually first or second, for our proofs to hold but we shall not belabor this point.

In the sequel, D is a simply connected domain of E^4 with subdomain σ having as its boundary the closed hypersurface $\partial\sigma$. Volume elements of σ are denoted dV while the (quaternion) oriented, outwardly directed surface elements of $\partial\sigma$ are denoted $d\mathbf{Q}$. Introducing the **quaternion gradient operator**

$$\square = \frac{\partial}{\partial w} + \nabla = \frac{\partial}{\partial w} + \mathbf{i} \frac{\partial}{\partial x} + \mathbf{j} \frac{\partial}{\partial y} + \mathbf{k} \frac{\partial}{\partial z},$$

we have the following useful result.

THEOREM 2.1. Let $\mathbf{F} = \phi + \psi$ be a function of the quaternion variable $\mathbf{q} = w + \mathbf{r}$, then

$$(1) \quad \int_{\partial\sigma} (d\mathbf{Q})\mathbf{F} = \int_{\sigma} \square \mathbf{F} dV.$$

Proof. Equation (1) is a quaternion form of the Gauss divergence theorem for four dimensions. Let $d\mathbf{Q} = dQ_0 + dQ_1 \mathbf{i} + dQ_2 \mathbf{j} + dQ_3 \mathbf{k}$. If M is the matrix

$$\begin{bmatrix} \phi & \psi_1 & \psi_2 & \psi_3 \\ -\psi_1 & \phi & -\psi_3 & \psi_2 \\ -\psi_2 & \psi_3 & \phi & -\psi_1 \\ -\psi_3 & -\psi_2 & \psi_1 & \phi \end{bmatrix}$$

and $[d\mathbf{q}] = (dQ_0, dQ_1, dQ_2, dQ_3)$ is a row vector having the same components as the quaternion $d\mathbf{Q}$, then, the matrix product $[dq]M$ is a row vector with the same components as the quaternion product $d\mathbf{Q}\mathbf{F}$. By the Gauss Theorem

$$\int_{\partial\sigma} [d\mathbf{q}]M = \int_{\sigma} \text{div}(M) dV,$$

where the matrix divergence of M is to be taken.

It is readily verified that $\text{div}(M)$ is a row vector whose four components are the same as those of the quaternion

$$\begin{aligned} \square \mathbf{F} &= \left(\frac{\partial}{\partial w} + \nabla \right) (\phi + \psi) \\ (2) \quad &= \frac{\partial \phi}{\partial w} + \nabla \phi + \frac{\partial \psi}{\partial w} - \nabla \cdot \psi + \nabla \times \psi, \end{aligned}$$

which establishes the result.

Similarly, we may demonstrate the alternate form of this result:

$$\int_{\partial\sigma} \mathbf{F} d\mathbf{Q} = \int_{\sigma} \mathbf{F} \square dV,$$

where the gradient operator is understood to act on the function F to its left.

3. Regular quaternion functions. In seeking to construct a differential and integral calculus of quaternion functions the first step would seem to be definition of a derivative. A (right) quaternion derivative of the function \mathbf{F} might be formed by requiring the limit

$$d\mathbf{F}/d\mathbf{q} = \lim (\mathbf{F}(\mathbf{q} + \Delta\mathbf{q}) - \mathbf{F}(\mathbf{q}))/\Delta\mathbf{q}$$

to exist as $\Delta\mathbf{q} \rightarrow 0$ and be independent of path for all increments $\Delta\mathbf{q}$. By considering four linearly independent increments Δw , $\Delta x\mathbf{i}$, $\Delta y\mathbf{j}$, $\Delta z\mathbf{k}$ one can derive a set of over-determined partial differential equations to be satisfied relating the components of \mathbf{F} under such conditions. This approach leads to nothing productive since, even for the simple function \mathbf{q}^2 , the ratio of $\Delta\mathbf{F}$ to $\Delta\mathbf{q}$ is not independent of $\Delta\mathbf{q}$, as was first observed by Hamilton [9]. The best one can do is to define scalar directional derivatives under the definition

$$d_{\mathbf{n}}\mathbf{F} = \lim (\mathbf{F}(\mathbf{q} + \varepsilon\mathbf{n}) - \mathbf{F}(\mathbf{q}))/\varepsilon$$

with ε real, $\varepsilon \rightarrow 0$, and \mathbf{n} a unit quaternion in the desired direction. The vector Taylor series expansion theorem in any direction can be then obtained but no real calculus results since only directionally dependent quantities are encountered. These ideas were first put forward by Hamilton himself in his *Elements of Quaternions*, [9].

To avoid the above difficulties, a weaker condition than path independence of

the differential ratio must be adopted. For a continuous function of the complex variable $z = x + iy$, the assertion

$$\int_C f(z)dz = 0$$

for every closed contour, C , in a domain of the z -plane is equivalent to the regularity of f in that domain (Morera's theorem). An alternate approach which suggests itself is the following.

A function \mathbf{F} of the quaternion variable \mathbf{q} is said to be **left regular** in D if

$$(3) \quad \int_{\partial\sigma} d\mathbf{Q}\mathbf{F} = 0$$

for every closed hypersurface, $\partial\sigma$, in D .

A **right regular** function is defined in similar manner by requiring the vanishing of the integral $\int_{\partial\sigma} \mathbf{F}(\mathbf{q})d\mathbf{Q}$ under the same circumstances. The following properties of regular functions are easily established.

LEMMA 3.1. *If $\mathbf{F}(\mathbf{q})$ is right (left) regular in D and \mathbf{q}_0 is a constant quaternion, then $\mathbf{F}(\mathbf{q} - \mathbf{q}_0)$ is also right (left) regular in D .*

LEMMA 3.2. *If \mathbf{F} is right regular in D and \mathbf{G} is left regular in D then $\int_{\partial\sigma} \mathbf{F}d\mathbf{Q}\mathbf{G} = 0$ for any closed hypersurface, $\partial\sigma$, in D .*

THEOREM 3.1. *The function $\mathbf{F} = \phi + \psi$ is left regular in D if and only if*

$$(4) \quad \frac{\partial\phi}{\partial w} = \nabla \cdot \psi$$

$$(5) \quad \nabla\phi = -\frac{\partial\psi}{\partial w} - \nabla \times \psi.$$

Proof. This result follows directly from (1) and (2) since (4) and (5) are equivalent to the single quaternion equation $\square\mathbf{F} = 0$.

The equations satisfied by the components of a right regular function are identical to (4) and (5) with the sign preceding the cross product in (5) changed to plus and the identical to the quaternion equation $\mathbf{F}\square = 0$. If a function is simultaneously left and right regular or, briefly, regular, then $\nabla \times \psi = 0$ and ψ is the gradient of a scalar potential function, $\psi = \nabla\Phi$. In this case, (4) and (5) are replaced by

$$\frac{\partial\phi}{\partial w} = \Delta\Phi, \quad \nabla\left(\phi + \frac{\partial\Phi}{\partial w}\right) = 0,$$

where Δ denotes the three dimensional Laplacian operator in x , y and z . These last two equations have some application in the study of fluid flow [3].

COROLLARY 3.1.1. *Each component of a left or right regular function satisfies Laplace's equation in the four variables w , x , y and z .*

Proof. Taking the divergence of both sides of (5) we obtain

$$\Delta\phi = -\nabla \cdot \frac{\partial\psi}{\partial w} = -\frac{\partial}{\partial w}(\nabla \cdot \psi).$$

From (4)

$$\Delta\phi = -\frac{\partial}{\partial w}\left(\frac{\partial\phi}{\partial w}\right) = -\frac{\partial^2\phi}{\partial w^2}.$$

Thus,

$$\Delta\phi + \frac{\partial^2\phi}{\partial w^2} = \Delta_4\phi = 0$$

as required for the scalar part of \mathbf{F} . From (5) we derive

$$\begin{aligned} -\frac{\partial^2\psi}{\partial w^2} &= \frac{\partial}{\partial w}(\nabla \times \psi) + \frac{\partial}{\partial w}\nabla\phi \\ &= -\nabla \times (\nabla\phi + \nabla \times \psi) + \nabla(\nabla \cdot \psi) = \Delta\psi, \end{aligned}$$

so that

$$\Delta\psi + \frac{\partial^2\psi}{\partial w^2} = 0.$$

As might be expected, given a scalar function ϕ sufficient differentiability, a vector function ψ can be found so that $\phi + \psi$ constitutes a regular function of \mathbf{q} , [1]. Due to the well-known maximum principle for solutions of Laplace's equation we have the following analogue of Liouville's theorem.

COROLLARY 3.1.2. *The only quaternion function regular with bounded norm in all E^4 is a constant.*

The concept of regularity may be extended to include functions regular in $\bar{\mathbf{q}}$.

DEFINITION. A function $\mathbf{F} = \phi + \psi$ is said to be **left regular** in $\bar{\mathbf{q}}$ for a domain D provided

$$\int_{\partial\sigma} d\bar{\mathbf{Q}}\mathbf{F} = 0$$

for every closed hypersurface, $\partial\sigma$, in D .

Right regularity in $\bar{\mathbf{q}}$ is defined in the obvious manner through the vanishing of $\int_{\partial\sigma} \mathbf{F}d\bar{\mathbf{Q}}$ in D . Necessary and sufficient conditions for \mathbf{F} to be left (right) regular in \mathbf{q} are $\bar{\square}\mathbf{F} = 0$ ($\mathbf{F}\bar{\square} = 0$) in D where

$$\bar{\square} = \frac{\partial}{\partial w} - \nabla.$$

Regular functions of $\bar{\mathbf{q}}$ also satisfy Corollaries 3.1.1 and 3.1.2. A function, \mathbf{F} , is left

(right) regular in $\bar{\mathbf{q}}$ only if its conjugate, $\bar{\mathbf{F}}$, is right (left) regular in \mathbf{q} . Further, the only functions simultaneously regular in both \mathbf{q} and $\bar{\mathbf{q}}$ are constants.

4. Generation of regular functions. Under the foregoing definitions, one hopes that a norm convergent quaternion power series of the form

$$\sum_{n=0}^{\infty} \mathbf{a}_n (\mathbf{q} - \mathbf{q}_0)^n,$$

where the \mathbf{a}_n are constant quaternions, would be a regular function of \mathbf{q} . Thus, for every regular function of the complex variable z one could generate an analogous regular function of \mathbf{q} by formally replacing z by \mathbf{q} in the power series expansion. It is the perversity of the quaternion calculus that even simple powers of q are not regular functions. For example, the scalar part of \mathbf{q}^2 is $w^2 - \mathbf{r} \cdot \mathbf{r}$ which does not satisfy Laplace's equation and hence cannot be regular in \mathbf{q} . Nevertheless, there is a close connection between convergent quaternion power series and regular functions. We shall term quaternion functions defined by norm convergent power series to be **analytic functions** and shall restrict ourselves to power series with real coefficients.

The formal device of replacing z by \mathbf{q} in a series expansion can be carried out in a more systematic manner. Let $f(z) = u(x, y) + iv(x, y)$ be a regular function of the complex variable $x + iy$ in some domain. We generate a quaternion function \mathbf{F} from f by replacing x with w , y with $r = (x^2 + y^2 + z^2)^{\frac{1}{2}}$ and i with $\mathbf{e}_r = \mathbf{r}/r$ so that

$$\mathbf{F}(\mathbf{q}) = u(w, r) + \mathbf{e}_r v(w, r).$$

Since z^n is replaced by \mathbf{q}^n this method yields the same result as the power series substitution. We inquire as to whether or not the function \mathbf{F} thus generated is regular in q . Instead of attempting to verify (4) and (5), we shall check the necessary conditions $\Delta_4(u + \mathbf{e}_r v) = 0$. We find that

$$\begin{aligned} \frac{\partial^2}{\partial w^2} (u + \mathbf{e}_r v) &= \frac{\partial^2 u}{\partial w^2} + \mathbf{e}_r \frac{\partial^2 v}{\partial w^2}, \\ \Delta(u + \mathbf{e}_r v) &= \frac{2}{r} \frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial r^2} + 2 \left(\frac{1}{r} \frac{\partial v}{\partial r} - \frac{1}{r^2} v \right) \mathbf{e}_r + \frac{\partial^2 v}{\partial r^2} \mathbf{e}_r. \end{aligned}$$

Since

$$\frac{\partial^2 u}{\partial w^2} + \frac{\partial^2 u}{\partial r^2} = \frac{\partial^2 v}{\partial w^2} + \frac{\partial^2 v}{\partial r^2} = 0,$$

then

$$(6) \quad \Delta_4(u + \mathbf{e}_r v) = \frac{2}{r} \frac{\partial u}{\partial r} + 2 \left(\frac{1}{r} \frac{\partial v}{\partial r} - \frac{1}{r^2} v \right) \mathbf{e}_r.$$

The only functions generated in this manner whose components satisfy Laplace's equation are constants or linear functions of q . Using $\partial u / \partial r = -\partial v / \partial w$, we may

rewrite (6) as

$$\Delta_4(u + \mathbf{e}_r v) = 2 \left[-\frac{\partial}{\partial w} \left(\frac{v}{r} \right) + \mathbf{e}_r \frac{\partial}{\partial r} \left(\frac{v}{r} \right) \right].$$

Since the special variables x, y, z only occur in the combination r , this result appears to be a special case of the more general equation

$$(7) \quad \Delta_4(u + \mathbf{e}_r v) = 2 \left(-\frac{\partial}{\partial w} + \nabla \right) \left(\frac{v}{r} \right).$$

Since

$$\Delta_4(\mathbf{u} + \mathbf{e}_r v) = \square \bar{\square} (u + \mathbf{e}_r v) = -2 \bar{\square} \left(\frac{v}{r} \right),$$

we deduce that

$$(8) \quad \square \mathbf{F} = -2 \frac{v}{r},$$

as may be readily verified. The equation (8) holds only for functions \mathbf{F} , constructed in the above manner. If \mathbf{F} is generated from a function regular in the complex variable $\bar{z} = x - iy$, the corresponding result obtained is

$$(9) \quad \bar{\square} \mathbf{F} = 2 \frac{v}{r}.$$

Equations (8) and (9) yield the relation

$$(10) \quad \bar{\square} \mathbf{F} = 2 \left(\frac{\partial \mathbf{F}}{\partial w} + \frac{v}{r} \right),$$

which may be applied if \mathbf{F} is generated from a regular function of z .

The symmetry of the generating process shows that the generated function must be regular (both left and right) if it is either left or right regular and, therefore, must satisfy $\square \mathbf{F} = 0$. Equation (8) shows that functions generated from regular functions are not regular; however,

$$\Delta_4 \left(\frac{v}{r} \right) = \frac{1}{r} \left(\frac{\partial^2 v}{\partial w^2} + \frac{\partial^2 v}{\partial r^2} \right) = 0.$$

We have proved the following:

THEOREM 4.1. *If \mathbf{F} is generated from a regular function f of z then the function $\Delta_4 \mathbf{F}$ is a regular function of \mathbf{q} .*

COROLLARY 4.1.1. *The norm convergent series $\sum \mathbf{a}_n \Delta_4 \mathbf{q}^n$ is regular in \mathbf{q} .*

COROLLARY 4.1.2. *Each component of a function \mathbf{F} generated as above satisfies the biharmonic equation $\Delta_4 \Delta_4 \mathbf{F} = 0$.*

COROLLARY 4.1.3. *Let v be harmonic in w and r . Then the quaternion function $\bar{\square}(v/r)$ is regular in \mathbf{q} .*

THEOREM 4.2. *Let \mathbf{F} be generated from the function f regular in z then*

$$(11) \quad \Delta_4 \mathbf{F} = \frac{2}{r} \mathbf{e}_r \left(\frac{\partial \mathbf{F}}{\partial w} - \frac{v(w, r)}{r} \right).$$

Proof. Applying the operator $\bar{\square}$ to both sides of (8), we find

$$(12) \quad \bar{\square} \square \mathbf{F} = \Delta_4 \mathbf{F} = -2 \bar{\square} \left(\frac{v}{r} \right) = -\frac{2}{r} \left(\bar{\square} v + \frac{v}{r} \mathbf{e}_r \right).$$

Since v may be written as $v(w, r) = \frac{1}{2}(\mathbf{e}_r \bar{\mathbf{F}} - \mathbf{e}_r \mathbf{F})$, we have

$$(13) \quad \bar{\square} v = \frac{1}{2} \bar{\square}(\mathbf{e}_r \bar{\mathbf{F}}) - \frac{1}{2} \bar{\square}(\mathbf{e}_r \mathbf{F}).$$

The function $\mathbf{e}_r \bar{\mathbf{F}}$ is generated from the function $if^{\bar{}}$ which is regular in \bar{z} while $\mathbf{e}_r \mathbf{F}$ is generated from if which is regular in z , so by (9) and (10), equation (13) becomes

$$\bar{\square} v = \frac{1}{2} \left(2 \frac{u}{r} \right) - \frac{1}{2} \left(2 \mathbf{e}_r \frac{\partial \mathbf{F}}{\partial w} + \frac{u}{r} \right) = -\mathbf{e}_r \frac{\partial \mathbf{F}}{\partial w}.$$

Equation (11) now follows from (12). Fueter's two formulas for $\Delta_4 \mathbf{q}^n$ and $\Delta_4 \mathbf{q}^{-n}$ [1, p. 316] are special cases of (11).

5. The Cauchy-Fueter integral formula. Cauchy's integral formula expresses the value of a regular function at a point interior to a closed contour in terms of the integral of its values on the contour. An analogous but more complicated theorem holds for regular functions of \mathbf{q} . We shall need the following fundamental theorem.

THEOREM 5.1. *Let $\partial\sigma$ be a closed hypersurface in E^4 containing the point \mathbf{q}_0 , then*

$$(14) \quad \int_{\partial\sigma} \Delta_4(\mathbf{q} - \mathbf{q}_0)^n d\mathbf{Q} = \begin{cases} 0 & n = 0, 1, \dots \\ 8\pi^2 & n = -1 \\ 0 & n = -2, -3, \dots \end{cases}$$

Proof. For n a non-negative integer $\Delta_4(\mathbf{q} - \mathbf{q}_0)^n$ is regular in E^4 and the result follows directly from the definition of regularity. If n is a negative integer the desired results can all be obtained from the case $n = -1$ by differentiation under the integral sign with respect to the scalar part of \mathbf{q}_0 . In fact, if

$$(15) \quad \int_{\partial\sigma} \Delta_4(\mathbf{q} - \mathbf{q}_0)^{-1} d\mathbf{Q} = 8\pi^2,$$

then

$$\begin{aligned} \frac{\partial^k}{\partial w_0^k} \int_{\partial\sigma} \Delta_4(\mathbf{q} - \mathbf{q}_0)^{-1} d\mathbf{Q} &= \int_{\partial\sigma} \Delta_4 \frac{\partial^k}{\partial w_0^k} (\mathbf{q} - \mathbf{q}_0)^{-1} d\mathbf{Q} \\ &= (k!) \int_{\partial\sigma} \Delta_4(\mathbf{q} - \mathbf{q}_0)^{-(1+k)} d\mathbf{Q} = 0, \end{aligned}$$

with $\mathbf{q}_0 = w_0 + \mathbf{r}_0$ and $k = 1, 2, \dots$.

All that remains is to prove (15). In view of Lemma 3.1 we need only to establish the case where $\mathbf{q}_0 = 0$ and $\partial\sigma$ is a hypersurface enclosing the point $\mathbf{q}_0 = 0$. Since $\Delta_4 \mathbf{q}^{-1}$ is regular except at $\mathbf{q} = 0$,

$$\int_{\partial\sigma} \Delta_4 \mathbf{q}^{-1} d\mathbf{Q} = - \int_{|\mathbf{q}|=1} \Delta_4 \mathbf{q}^{-1} d\mathbf{Q}.$$

Equation (7) can be used to find $\Delta_4 \mathbf{q}^{-1}$. Because $v(w, r) = -r/\rho^2$ where $\rho^2 = w^2 + r^2$, then $\Delta_4 \mathbf{q}^{-1} = -(4/\rho^2) \mathbf{q}^{-1}$. The scalar surface element of a sphere having radius $|\mathbf{q}|$ in E^4 is $|\mathbf{q}|^3 dS$, where dS is the surface element of the corresponding unit sphere in E^4 , [10, p. 677]. The oriented surface element for a sphere of radius $|\mathbf{q}|$ is therefore

$$(16) \quad d\mathbf{Q} = |\mathbf{q}|^2 \mathbf{q} dS.$$

The integral in question becomes

$$- \int_{|\mathbf{q}|=1} \Delta_4 \mathbf{q}^{-1} d\mathbf{Q} = 4 \int_{|\mathbf{q}|=1} \mathbf{q}^{-1} \mathbf{q} dS = 8\pi^2,$$

since the surface area of the unit sphere in E^4 is $2\pi^2$, [10, p. 677].

The previous theorem leads one to expect that the functions $\Delta_4 \mathbf{q}^n$ will play roughly the same role in the quaternion calculus that the functions z^n play in ordinary complex analysis. Given a function \mathbf{F} defined by a Laurent type series

$$\mathbf{F}(\mathbf{q}) = \sum_{n=-\infty}^{\infty} \mathbf{a}_n (\mathbf{q} - \mathbf{q}_0)^n$$

we deduce formally

$$\Delta_4 \mathbf{F}(\mathbf{q}) (\mathbf{q} - \mathbf{q}_0)^{-k} = \sum_{n=-\infty}^{\infty} \mathbf{a}_n \Delta_4 (\mathbf{q} - \mathbf{q}_0)^{n-k},$$

from which we derive

$$\mathbf{a}_{k-1} = \mathbf{F}(\mathbf{q}_0) = \frac{1}{8\pi^2} \int_{\partial\sigma} \Delta_4 \mathbf{F}(\mathbf{q}) (\mathbf{q} - \mathbf{q}_0)^{-k} d\mathbf{Q}.$$

Of more interest is the following analogue of the Cauchy integral formula.

THEOREM 5.2. *Let \mathbf{F} be a regular function of \mathbf{q} in D . If $\partial\sigma$ is a hypersurface in D containing the point \mathbf{q}_0 , then*

$$\mathbf{F}(\mathbf{q}_0) = \frac{1}{8\pi^2} \int_{\partial\sigma} \mathbf{F}(\mathbf{q}) d\mathbf{Q} \Delta_4(\mathbf{q} - \mathbf{q}_0)^{-1}.$$

Proof. For ε small enough, the hypersurface centered at \mathbf{q}_0 defined by $|\mathbf{q} - \mathbf{q}_0| = \varepsilon$ lies inside $\partial\sigma$. In the region between the surface of the ε -sphere and $\partial\sigma$ both \mathbf{F} and $\Delta_4(\mathbf{q} - \mathbf{q}_0)^{-1}$ are regular so that, using Lemma 3.2 we can show that

$$\frac{1}{8\pi^2} \int_{\partial\sigma} \mathbf{F}(\mathbf{q}) d\mathbf{Q} \Delta_4(\mathbf{q} - \mathbf{q}_0)^{-1} = \frac{1}{8\pi^2} \int_{|\mathbf{q} - \mathbf{q}_0| = \varepsilon} \mathbf{F}(\mathbf{q}) d\mathbf{Q} \Delta_4(\mathbf{q} - \mathbf{q}_0)^{-1}.$$

The surface element for the last integral is found by replacing \mathbf{q} with $\mathbf{q} - \mathbf{q}_0$ in (16); thus,

$$d\mathbf{Q} = |\mathbf{q} - \mathbf{q}_0|^2 (\mathbf{q} - \mathbf{q}_0) dS.$$

The function $\mathbf{F}(\mathbf{q})$ is to have sufficient differentiability so that

$$\mathbf{F}(\mathbf{q}) = \mathbf{F}(\mathbf{q}_0) + O(|\mathbf{q} - \mathbf{q}_0|), \quad |\mathbf{q} - \mathbf{q}_0| \rightarrow 0.$$

The limit of the last integral as $\varepsilon \rightarrow 0$ is therefore found to be

$$\begin{aligned} \lim \frac{4}{8\pi^2} \int_{|\mathbf{q} - \mathbf{q}_0| = 1} \mathbf{F}(\mathbf{q}) \varepsilon^2 (\mathbf{q} - \mathbf{q}_0) \varepsilon^{-2} (\mathbf{q} - \mathbf{q}_0)^{-1} dS \\ = \lim \frac{1}{2\pi^2} \int_{|\mathbf{q} - \mathbf{q}_0| = 1} (\mathbf{F}(\mathbf{q}_0) + O(\varepsilon)) dS = \mathbf{F}(\mathbf{q}_0) \end{aligned}$$

as required.

It is essential in Theorem 5.2 that the terms in the integral be separated by the differential $d\mathbf{Q}$ since $\mathbf{F} \cdot \Delta_4(\mathbf{q} - \mathbf{q}_0)^{-1}$ is not generally regular even if \mathbf{F} is. Many properties of regular functions such as the existence of series expansions, mean value theorems, etc., can be proved from (17) in much the same manner as is done in complex analysis. We give only one such example, the familiar Poisson integral formula for $n = 4$, [10, p. 265].

COROLLARY 5.2.1. *Let $\mathbf{F} = \phi + \psi$ be regular in \mathbf{q} for $|\mathbf{q}| \leq \rho$ and let $\mathbf{q}_0 = \mathbf{w}_0 + \mathbf{r}_0$ be a point such that $|\mathbf{q}_0| = R$, where $\rho > R$, then*

$$(18) \quad \phi(\mathbf{q}_0) = \frac{\rho^2(\rho^2 - R^2)}{2\pi^2} \int_{|\mathbf{q} - \mathbf{q}_0| = 1} \frac{\phi(\mathbf{q}) dS}{(\rho^2 + R^2 - 2\rho R \cos(\theta))^2}$$

where $\cos(\theta)$ is defined by

$$(19) \quad \cos(\theta) = (\mathbf{w}\mathbf{w}_0 + \mathbf{r} \cdot \mathbf{r}_0) / |\mathbf{q}| |\mathbf{q}_0|.$$

Proof. From (19)

$$\begin{aligned} |\mathbf{q}| |\mathbf{q}_0| \cos(\theta) &= \mathbf{w}\mathbf{w}_0 + \mathbf{r} \cdot \mathbf{r}_0 \\ &= |\mathbf{q}|^2 + |\mathbf{q}_0|^2 - |\mathbf{q} - \mathbf{q}_0|^2 \end{aligned}$$

so that (18) becomes

$$(20) \quad \phi(\mathbf{q}_0) = \frac{\rho^2(\rho^2 - R^2)}{2\pi^2} \int_{|\mathbf{q}-\mathbf{q}_0|=1} \frac{\phi(\mathbf{q})dS}{|\mathbf{q}-\mathbf{q}_0|^4}$$

From Theorem 5.2,

$$(21) \quad \begin{aligned} \mathbf{F}(\mathbf{q}_0) &= \frac{1}{8\pi^2} \int_{|\mathbf{q}-\mathbf{q}_0|=\rho} \mathbf{F}(\mathbf{q})d\mathbf{Q}\Delta_4(\mathbf{q}-\mathbf{q}_0)^{-1} \\ &= \frac{1}{2\pi^2} \int_{|\mathbf{q}-\mathbf{q}_0|=1} \frac{\mathbf{F}(\mathbf{q})\rho^2\mathbf{q}dS(\mathbf{q}-\mathbf{q}_0)^{-1}}{|\mathbf{q}-\mathbf{q}_0|^2} \end{aligned}$$

If $\mathbf{q}_0^* = \rho^2\bar{\mathbf{q}}_0^{-1}$ then $|\mathbf{q}_0^*| > \rho^2/R > \rho$ and

$$(22) \quad \begin{aligned} 0 &= \frac{1}{8\pi^2} \int_{|\mathbf{q}-\mathbf{q}_0|=\rho} \mathbf{F}(\mathbf{q})d\mathbf{Q}(\mathbf{q}-\mathbf{q}_0)^{-1} \\ &= \frac{1}{2\pi^2} \int_{|\mathbf{q}-\mathbf{q}_0|=1} \left(\frac{R}{\rho}\right)^2 \frac{\rho^2\mathbf{q}dS\mathbf{q}^{-1}(\bar{\mathbf{q}}_0-\bar{\mathbf{q}})^{-1}\bar{\mathbf{q}}_0}{|\mathbf{q}-\mathbf{q}_0|^2}. \end{aligned}$$

Dividing out the constant $(R/\rho)^2$ from (22) and subtracting (21) from (22) we find

$$\begin{aligned} \mathbf{F}(\mathbf{q}_0) &= \frac{1}{2\pi^2} \int_{|\mathbf{q}-\mathbf{q}_0|=1} \frac{\mathbf{F}(\mathbf{q})\rho^2}{|\mathbf{q}-\mathbf{q}_0|^2} [\mathbf{q}(\mathbf{q}-\mathbf{q}_0)^{-1} + (\bar{\mathbf{q}}-\bar{\mathbf{q}}_0)^{-1}\bar{\mathbf{q}}_0]dS \\ &= \frac{\rho^2 - R^2}{2\pi^2} \int_{|\mathbf{q}-\mathbf{q}_0|=1} \frac{\mathbf{F}(\mathbf{q})dS}{|\mathbf{q}-\mathbf{q}_0|^4} \end{aligned}$$

which proves the equation (20) and hence the theorem when the scalar parts of the last equation are equated.

6. Applications. Aside from older mechanics texts which sometimes treat rotations in the quaternion form, they are seldom encountered except with their cousins octernions and Clifford numbers in the factorization of relativistic energy equations, [11]. Instead of studying Laplace's equation in 4 variables, one generally wants to consider the wave operator,

$$\Delta - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}.$$

Formal replacement of w to ict changes one equation into the other. The equations for right regular quaternion functions then become

$$(23) \quad -\frac{i}{C} \frac{\partial \phi}{\partial t} = \nabla \cdot \psi,$$

$$(24) \quad \nabla \phi = \frac{i}{C} \frac{\partial \psi}{\partial t} + \nabla \times \psi.$$

The resemblance of (23) to a conservation equation suggests the further substitution $\phi = -i\lambda$ to obtain from (23) and (24)

$$(25) \quad \frac{1}{c} \frac{\partial \lambda}{\partial t} + \nabla \cdot \psi = 0,$$

$$(26) \quad \nabla \lambda = -\frac{1}{c} \frac{\partial \psi}{\partial t} + i \nabla \times \psi.$$

We expect λ to be real and ψ to have real components. Equations (25) and (26) describe a variety of physical systems. If we identify $\lambda = ce$ and $\psi = c^2\pi$ where c is the speed of light *in vacuo*, e and π are the relativistic energy and momentum densities, respectively, of the system under consideration, then, (25) and (26) break into the three equations

$$\frac{\partial e}{\partial t} + \nabla \cdot (c^2\pi) = 0$$

$$\frac{1}{c} \frac{\partial \pi}{\partial t} + \nabla \left(\frac{e}{c} \right) = 0$$

$$\nabla \times \pi = 0.$$

These three equations, the first of which is the conservation of mass-energy, constitute the basis of relativistic mechanics in the absence of electromagnetic forces, [11, p. 272]. Defining the relativistic momentum quaternion $\mathbf{P} = -ie + c\pi$ and the operator

$$\square^* = \frac{-i}{c} \frac{\partial}{\partial t} + \nabla$$

we have the following quaternion expression for these equations:

$$\square^* \mathbf{P} = 0.$$

Thus, we have proved the following result.

THEOREM 6.1. *In the absence of electromagnetic forces, the momentum quaternion, \mathbf{P} , is a (formal) regular function of the quaternion variable $ict + r$.*

Maxwell's equations

$$\nabla \cdot \mathbf{H} = 0, \quad \nabla \cdot \mathbf{E} = \rho$$

$$\frac{1}{c} \frac{\partial \mathbf{H}}{\partial t} + \nabla \times \mathbf{E} = 0$$

$$\frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} + \mathbf{J} = \nabla \times \mathbf{H}$$

are likewise expressed in the simple quaternion form

$$(27) \quad \square^*(\mathbf{E} + i\mathbf{H}) = -\rho + \frac{i}{c} \mathbf{J}.$$

If ϕ and \mathbf{A} are the usual Hertzian scalar and vector potentials for \mathbf{E} and \mathbf{H} , [12, p. 212], the electromagnetic field is derivable from the quaternion potential function $i\phi + \mathbf{A}$ through the equation

$$(28) \quad \bar{\square}^*(i\phi + \mathbf{A}) = i(\mathbf{E} + i\mathbf{H}).$$

From (27) and (28) we obtain

$$(29) \quad \square^* \bar{\square}^*(i\phi + \mathbf{A}) = \left(-\frac{1}{c^2} \frac{\partial^2}{\partial t^2} + \Delta \right) (i\phi + \mathbf{A}).$$

In component form (29) yield the equations relating the electromagnetic potential with the charge and current densities

$$\square^2 \phi = \rho, \quad \square^2 \mathbf{A} = \frac{1}{c} \mathbf{J},$$

where

$$\square^2 = \Delta - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}.$$

Acknowledgments. The author wishes to thank John Castelluci for translating portions of the original German, and the referee for his many valuable suggestions and improvements in the paper.

References

1. R. Fueter, Die Funktionentheorie der Differentialgleichungen $\Delta u = 0$ und $\Delta \Delta u = 0$ mit vier reellen Variablen, *Comment. Math. Helv.*, 7(1935) 307–30.
2. H. Haefeli, Hyperkomplexe Differentiale, *Comment. Math., Helv.* 20 (1947) 382–420.
3. A. Rose, On the use of a complex (quaternion) velocity potential in three dimensions, *Comment. Math. Helv.*, 24 (1950) 135–48.
4. R. Fueter, Über vierfachperiodische Funktionen, *Montsh. Math. Phys.*, 48 (1939) 161–69.
5. E. Bell, *Development of Mathematics*, McGraw-Hill, New York, 1945 Chapter IX.
6. M. Crowe, *A History of Vector Analysis*, Notre Dame Press, 1967.
7. C. Curtis, *MAA Studies in Modern Algebra*, Vol. II, Math. Assoc. of America, 1963, pgs. 108–11.
8. P. Duval, *Homographies, Quaternions, and Rotations*, Oxford Math. Monographs, Oxford Press, 1964.
9. W. Hamilton, *Elements of Quaternions*, Vol. I, Chapter II, 1860, reprinted by Chelsea, New York.
10. R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. II, Interscience, Wiley, New York, 1965.
11. A. Kyrala, *Theoretical Physics*, Sanders, Philadelphia, 1967.
12. H. Phillips, *Vector Analysis*, Wiley, New York, 1963.

INEQUALITIES FOR SUMS OF DISTANCES

G. D. CHAKERIAN, University of California at Davis and
M. S. KLAMKIN, Ford Motor Company

1. Introduction. Any triangle inscribed in a circle of radius R and having the center as an interior point has perimeter greater than $4R$. This theorem is discussed in [1], where the reader can find a number of references to papers dealing with generalizations and applications of the result. For example, it follows readily from this that any closed plane curve of length L can be covered by a circular disk of radius $L/4$. Assuming our circle is the unit circle centered at the origin and V_1, V_2, V_3 are the unit vectors representing the vertices of the triangle, the theorem becomes,

$$(1) \quad |V_1 - V_2| + |V_2 - V_3| + |V_1 - V_3| > 4.$$

Thus, we see that inequality (1) is satisfied by any three unit vectors whose convex hull has the origin as an interior point.

Our main purpose is to consider extensions of (1) to the case of more than three vectors and to higher dimensional spaces. Indeed, in Section 2, we shall give a simple proof of the following theorem for r unit vectors in Euclidean n -space E^n .

THEOREM 1. *If V_1, V_2, \dots, V_r are unit vectors in E^n and the origin is interior to their convex hull, then*

$$\sum |V_i - V_j| > 2(r - 1),$$

where the sum is over $1 \leq i < j \leq r$.

The following corollary, which is a special case of Conjecture 4 raised in [1], is an immediate consequence.

COROLLARY 1. *The total edge length of a simplex inscribed in a unit sphere in E^n with the center in its interior is greater than $2n$.*

We shall see that Theorem 1 follows from an even stronger inequality, namely,

$$(2) \quad \sum |V_i - V_j|^2 > 4(r - 1).$$

In Section 3, we give a second proof of Theorem 1. Although not as simple as the proof in Section 2, the idea of this proof leads to other interesting results and new proofs of some theorems discussed in later sections. Finally, we discuss some questions about upper bounds in Section 7.

2. Proof of Theorem 1. The proof depends on two known results. First we have the following easily established identity for r unit vectors in E^n .

$$(3) \quad \sum |V_i - V_j|^2 = r^2 - |\sum V_i|^2,$$

where as usual the sum on the left hand side is over $1 \leq i < j \leq r$ while that on the

right hand side is for $1 \leq i \leq r$. This identity, in fact, gives a solution to a problem on the 1968 William Lowell Putnam Mathematical Competition (see [2] where a derivation of (3) is given). Next, we use an inequality proved in [3], that is, with the conditions given in Theorem 1,

$$(4) \quad |\sum V_i| < r - 2.$$

For the sake of completeness and because of the relative inaccessibility of reference [3], we now give a proof of (4).

Since the V_i 's do not all lie in any half-space, the convex cone which is positively spanned by the vectors must be E^n . For if it was not, there would then exist some hyperplane [4] such that the cone would lie entirely to one side of it. But this would violate the hypothesis. Consequently, there exists a set of numbers a_1, \dots, a_r where $a_i \geq 0$, $i = 1, \dots, r$, $\sum a_i = 1$ such that

$$a_1 V_1 + \dots + a_r V_r = 0.$$

We now show that $a_i \leq 1/2$. Assume the contrary that $a_i > 1/2$. Then

$$|a_2 V_2 + \dots + a_r V_r| = |-a_1 V_1|.$$

The r.h.s. $> 1/2$ and the l.h.s. $\leq a_2 + \dots + a_r = 1 - a_1 < 1/2$ which is a contradiction. Finally, since

$$V_1 + \dots + V_r = (1 - 2a_1)V_1 + \dots + (1 - 2a_r)V_r,$$

$$|V_1 + \dots + V_r| < (1 - 2a_1) + \dots + (1 - 2a_r) = r - 2.$$

This result is best possible since one can get arbitrarily close to $r - 2$ by considering a sequence of convex polytopes converging to the degenerate case of a segment.

Inequality (2) is now immediate consequence of (4) in conjunction with (3). Since $|V_i - V_k| \leq 2$, we have $|V_i - V_j|^2 \leq 2|V_i - V_j|$, so Theorem 1 follows from (2). This completes the proof.

The inequality sign in Theorem 1 can be replaced with equality if and only if one of the vectors is the negative of the remaining $r - 1$ vectors. This is easily deduced once one observes that in case of equality, we must have $|V_i - V_j|^2 = 2|V_i - V_j|$ for $1 \leq i < j \leq r$, so $|V_i - V_j|$ must always be 0 or 2. It should be noted that the inequality sign in (2) may be replaced by equality in other cases, e.g., when two vectors are antipodal and the remaining $r - 2$ coincide.

3. Another proof of Theorem 1. This proof involves an application of a formula from integral geometry, expressing the length of a curve on the unit sphere in terms of the expected number of intersections with a random great circle. To state the higher dimensional version of this, let S^{n-1} be the unit sphere in E^n , $n \geq 3$, and for each $u \in S^{n-1}$ let $G(u)$ be the great $(n - 2)$ -sphere orthogonal to u . In other words, $G(u)$ is the intersection of S^{n-1} with the $(n - 1)$ -dimensional subspace

orthogonal to u . Let du represent the normalized element of $(n-1)$ -dimensional rotation invariant surface measure on S^{n-1} , so that

$$\int_{S^{n-1}} du = 1.$$

Now suppose C is a finite collection of arcs of great circles on S^{n-1} of total length $L(C)$. For each $u \in S^{n-1}$, let $N(C, u)$ be the cardinality of $G(u) \cap C$. Then,

$$(5) \quad L(C) = \pi \int N(C, u) du,$$

where the integration is over all of S^{n-1} . A proof of this formula, in case $n = 3$, is found in [5]. The formula is more generally valid for rectifiable arcs. For a discussion of the proof in this case, at least for $n = 3$, see [6] and the references given therein. Because of its intrinsic interest we now sketch a proof of (5), using a method different from that used in [5].

If C_1, \dots, C_k are arcs of great circles, and $C = C_1 \cup C_2 \cup \dots \cup C_k$, then

$$N(C, u) = \sum_{i=1}^k N(C_i, u),$$

except for an exceptional set of directions constituting a set of zero spherical surface measure. Since also

$$L(C) = \sum_{i=1}^k L(C_i),$$

it follows that it suffices to prove (5) in case C is a single arc of a great circle. But in this case, the right hand side of (5) is a function $I(\sigma)$ depending only on the length σ of the arc C and not depending on its position. Moreover, I is an additive function in the sense that

$$I(\sigma_1 + \sigma_2) = I(\sigma_1) + I(\sigma_2),$$

as can be seen by subdividing the arc C . Since I can be shown to be a continuous function of σ , and since it is known that any additive continuous function $I(\sigma)$ must be of the form $I(\sigma) = \lambda\sigma$, for some constant λ , we have that I is proportional to the arclength of C . To calculate the proportionality constant λ , we compute I for a great semicircular arc C :

$$\lambda\pi = I(\pi) = \pi \int N(C, u) du = \pi,$$

since $N(C, u) = 1$ in this case, with a set of exceptions of zero measure. Formula (5) then follows.

Coming back to the proof of Theorem 1, let $P, Q \in S^{n-1}$, we let σ be the length

of the shortest great circle arc joining them and $d = |P - Q|$. We have $0 \leq \sigma \leq \pi$, hence

$$\frac{d}{2} = \sin \frac{\sigma}{2} \geq \frac{2}{\pi} \left(\frac{\sigma}{2} \right),$$

by Jordan's inequality [6, p. 33], in other words,

$$(6) \quad d \geq \frac{2}{\pi} \sigma.$$

Now suppose $V_1, \dots, V_r \in S^{n-1}$. For each $i < j$, set $d_{ij} = |V_i - V_j|$ and let σ_{ij} be the length of a shortest great circle arc joining V_i to V_j . Assuming the origin is interior to the convex hull of V_1, \dots, V_r , each hyperplane through the origin not containing any V_i must separate some k of the points, $1 \leq k \leq r - 1$, from the remaining $r - k$. Hence such a hyperplane crosses $k(r - k) \geq r - 1$ of the line segments determined by V_1, \dots, V_r . Let C be the collection of great circle arcs on S^{n-1} obtained by projecting the set of line segments with endpoints among V_1, \dots, V_r radially into the sphere. (In case a diameter is encountered, simply associate with it any semicircle with the same endpoints.) With $N(C, u)$ defined as above, the preceding argument shows that

$$(7) \quad N(C, u) \geq r - 1,$$

because $G(u)$ intersects C in the same number of points that the hyperplane through the origin orthogonal to u intersects the set of line segments determined by V_1, \dots, V_r . This is true at least for those u such that $G(u)$ contains no V_i , a set of measure 1. From (6) we have

$$(8) \quad \sum |V_i - V_j| = \sum d_{ij} \geq \frac{2}{\pi} \sum \sigma_{ij},$$

and from (5) and (7),

$$(9) \quad \sum \sigma_{ij} = \pi \int N(C, u) du \geq \pi (r - 1).$$

Theorem 1 now follows from (8) and (9). We must actually have strict inequality in (8) under the given hypothesis since not all σ_{ij} equal π .

Note that (8) is applicable also in case $n = 2$. If $r = 3$, we have in this case $\sum \sigma_{ij} = 2\pi$, obtaining still another proof of (1).

4. The smallest ball containing a space curve. The methods used in section 3 provide another way to prove that any closed curve of length L in E^n is contained in a ball of radius $L/4$. For other proofs of this, and related problems, see [8, 9] and references given therein.

It is not difficult to see that it suffices to prove the following theorem.

THEOREM 2. *Let K be a closed curve of length $L(K)$ contained in the unit ball in E^n . If K intersects S^{n-1} , the boundary of the ball, in a set whose convex hull contains the origin, then $L(K) \geq 4$.*

Proof. There exist a finite number r of points in $K \cap S^{n-1}$ whose convex hull contains the origin. We may label these points V_1, \dots, V_r in such a way that they lie cyclically on K , so that

$$(10) \quad L(K) \geq \sum |V_i - V_{i+1}|,$$

where the sum is over $1 \leq i \leq r$, and by convention $V_{r+1} \equiv V_1$. Let C be the closed curve on S^{n-1} obtained by projecting the polygon P with successive vertices V_1, \dots, V_r radially into the sphere (again with the convention that if $V_i - V_{i+1}$ is a diameter, then we associate with it a semicircle with the same endpoints). Every hyperplane through the origin intersects the polygon P in at least two points; hence every great $(n-2)$ -sphere $G(u)$ intersects C in at least two points. Thus, with the notation of Section 3, we have $N(C, u) \geq 2$, so from (5),

$$(11) \quad L(C) \geq 2\pi.$$

If we now apply (6), with $d_i = |V_i - V_{i+1}|$ and σ_i equal to the length of the radial projection of the segment $V_i V_{i+1}$ into the sphere, we obtain

$$(12) \quad L(K) \geq \sum d_i \geq \frac{2}{\pi} \sum \sigma_i = \frac{2}{\pi} L(C) \geq 4.$$

5. Generalization to points on a convex curve. A generalization of (1) in another direction is given in [1]. Let K be a plane convex curve containing the origin in its interior. Let V_1, V_2, V_3 be points on K whose convex hull has the origin as an interior point. Then it is shown in [1] that

$$(13) \quad |V_1 - V_2| + |V_2 - V_3| + |V_1 - V_3| > 2m,$$

where m is the length of the minimum chord of K passing through the origin. A simple inductive argument will give the following extension of (13).

THEOREM 3. *Let K be as above and let V_1, \dots, V_r be points on K whose convex hull contains the origin. Then if m is the minimum chord of K containing the origin, we have*

$$\sum |V_i - V_j| \geq m(r-1).$$

Proof. The case $r = 3$ is proved in [1]. Assume the result known for $r \leq k$, and suppose V_1, \dots, V_{k+1} are $k+1$ points on K with the origin in the interior of their convex hull. The convex hull of some k of the points, say V_1, \dots, V_k , contains the origin, so

$$(14) \quad \sum_{1 \leq i < j \leq k} |V_i - V_j| \geq m(k-1).$$

Now some triangle with one vertex at V_{k+1} and its other two vertices among V_1, \dots, V_k contains the origin. On the basis of the case $r = 3$ of our inequality, it is easy to show that the sum of the lengths of each pair of sides of this triangle is at least m . Thus $|V_{k+1} - V_\alpha| + |V_{k+1} - V_\beta| \geq m$ for some $\alpha, \beta \in \{1, \dots, k\}$, $\alpha \neq \beta$. In conjunction with (14), this implies $\sum |V_i - V_j| \geq mk$, where the sum is over $1 \leq i < j \leq k + 1$. By induction, this completes the proof.

REMARK. A close examination of the proof, keeping (13) in mind, shows that equality can hold in Theorem 3 only in the degenerate case when V_1, \dots, V_r lie on a minimum chord and one of the points is the negative of the other $r - 1$ points.

6. Extensions to Minkowski planes. Let K be a centrally symmetric convex curve centered at the origin in E^2 . For any $X \in E^2$ define the K -norm of X by

$$(15) \quad |X|_K = |X|/r,$$

where $|X|$ is the ordinary Euclidean norm of X and r is the (Euclidean) length of the radius of K in the direction of X . Then $| \cdot |_K$ is a norm for a Minkowski plane, or a two dimensional Banach space, whose unit circle is K .

Bearing a close relationship to the matters discussed in this article, there is a famous theorem of Fenchel stating that every closed curve has total curvature at least 2π (see [5], where a proof based on (5) is given). Indeed, the main step of the proof is embodied in our inequality (11). In the course of obtaining a version of Fenchel's theorem valid for Minkowski spaces, Laugwitz [10] showed that if V_1, V_2, V_3 are unit vectors in a Minkowski plane (i.e., $|V_i|_K = 1$, $i = 1, 2, 3$), with the origin interior to their convex hull, then

$$(16) \quad |V_1 - V_2|_K + |V_2 - V_3|_K + |V_1 - V_3|_K \geq 4.$$

We give another proof of (16) based on an idea from Section 3. This requires the following analogue of Jordan's inequality for Minkowski planes.

LEMMA. Let K be the unit circle in a Minkowski plane and $P, Q \in K$. Let $d = |P - Q|_K$, the Minkowski length of the chord joining P to Q . Let σ be the length of the smaller arc of K from P to Q , and let λ be half the perimeter of K , both measured in the given Minkowski metric. Then,

$$d \geq 2\sigma/\lambda.$$

Proof. Let P' and Q' be the endpoints of the diameter of K parallel to the chord joining P to Q (with P', P, Q, Q' lying in that order along K). Let R be the point of intersection of the lines determined by PP' and QQ' respectively, assuming these lines are not parallel. Then R lies on the same side of the line through $P'Q'$ as P and Q , and the similarity transformation with R as fixed point and sending P to P' and Q to Q' also sends the arc σ to a convex arc σ' with endpoints P' and Q' . This transformation increases lengths by the factor $2/d$, so $\sigma' = 2\sigma/d$, where for con-

venience we use the same symbol to denote an arc and its length in the Minkowski metric. It is easy to see that the arc σ' lies inside the unit disk K . Using the fact that if K_1 is a closed convex curve inside a closed convex curve K_2 , then K_1 is not longer than K_2 (proved in a Minkowski plane in exactly the same way as in the Euclidean plane), one obtains that σ' has length at most half K . That is, $2\sigma/d = \sigma' \leq \lambda$, as we wanted to prove. If the lines through PP' and QQ' are parallel, the proof proceeds the same way, with the similarity transformation replaced by a translation.

To prove (16), let $d_i = |V_i - V_{i+1}|_K$, $i = 1, 2, 3$, with $V_4 \equiv V_1$, and σ_i equal to the length of the shorter arc of K from V_i to V_{i+1} . By the lemma, we have $d_i \geq 2\sigma_i/\lambda$, $i = 1, 2, 3$. Hence

$$\sum |V_i - V_{i+1}|_K = \sum d_i \geq \frac{2}{\lambda} \sum \sigma_i = 4,$$

which proves the inequality (16).

The same inductive argument used to establish Theorem 3 can be used in conjunction with (16) to prove the following generalization.

THEOREM 4. *Let V_1, V_2, \dots, V_r be unit vectors whose convex hull contains the origin in a given Minkowski plane. Then*

$$\sum |V_i - V_j|_K \geq 2(r-1).$$

Inequality (16) asserts that any triangle inscribed in the unit circle K in a Minkowski plane and having the origin in its interior has perimeter at least 4. This can be trivially extended as follows: *If P is an interior point of K , then any triangle inscribed in K with P in its interior has perimeter at least twice the minimum chord of K through P .* To prove this, observe that if the triangle contains the origin, then the result follows from (16). If the triangle, with vertices V_1, V_2, V_3 , does not contain the origin in its interior or on its boundary, then one side, say that through V_1 and V_2 , separates P from the origin. Then the chord of K through P parallel to $V_1 V_2$ has length $m \leq |V_1 - V_2|_K$, so $|V_1 - V_2|_K + |V_2 - V_3|_K + |V_1 - V_3|_K \geq 2|V_1 - V_2|_K \geq 2m$, from which the result follows.

It is natural to conjecture that the analogue of (13), for general convex curves in a Minkowski plane, is valid (the argument above shows it for a Minkowski circle). In that case, we would also have that the analogue of Theorem 3 is true in any Minkowski plane.

One might also conjecture that Theorem 1 holds in any Minkowski space. We are unable to prove this, although it is true that inequality (4) is valid in a Minkowski space.

7. Upper bounds. If V_1, V_2, \dots, V_r are any unit vectors in E^n , then it follows from (3) that

$$(17) \quad \sum |V_i - V_j|^2 \leq r^2,$$

with equality if and only if $\sum V_i = 0$. This, in fact, is the Putnam Examination problem (which was posed in case $n = 3$) mentioned in Section 2. We obtain from a straightforward application of the Cauchy inequality and (17), that

$$(18) \quad \sum |V_i - V_j| \leq \left(\frac{r^3(r-1)}{2} \right)^{\frac{1}{2}}.$$

Equality can hold in (18) only if the distances $|V_i - V_j|$, $i < j$, are all equal and $\sum V_i = 0$. Since this condition cannot be satisfied if $r > n + 1$ the inequality is not sharp in general. However, for $r = n + 1$, we obtain

$$\sum |V_i - V_j| \leq \left\{ \frac{n(n+1)^3}{2} \right\}^{\frac{1}{2}}$$

with equality if and only if the V_i are the vertices of a regular simplex inscribed in S^{n-1} . This is equivalent to the fact that of all simplices inscribed in the unit sphere, the regular simplex has maximum total edge length.

It is proved in [11, p. 155] that if V_1, \dots, V_r are unit vectors in E^2 then

$$\sum |V_i - V_j| \leq r \cot \frac{\pi}{2r},$$

with equality only when the vectors are the vertices of a regular r -gon. It is also shown there that $\sum |V_i - V_j|^{-1}$ takes its minimum value in the case of a regular r -gon. What the corresponding higher dimensional results should be (for $r > n + 1$) appears to be unknown. Analogous upper bounds for such sums of distances between unit vectors in Minkowski spaces also appear to be unknown, even in the case of two dimensions.

Note: The problem of obtaining an upper bound for the sum of all distances determined by n points on a unit ball has been studied more than Section 7 indicates; besides Fejes Tóth, other workers include R. Alexander, G. Björck, E. Hille, J. B. Kelly, F. Nielson, G. Pólya and G. Szegő, G. Sperling, K. B. Stolarsky, and H. S. Witsenhausen. A systematic treatment together with generalizations and extensive bibliography now exists, namely *Extremal Problems of Distance Geometry Related to Energy Integrals* by R. Alexander and K. B. Stolarsky (preprints available). For the unit sphere in the Euclidean space E^m with $m \geq 5$, the best results so far are in *Sums of Distances Between Points on a Sphere II*, K. B. Stolarsky, Proc. Amer. Math. Soc., to appear. For $m = 5$ the upper bound $S(n)$ is shown to satisfy

$$c_1 n^2 - c_2 n^{\frac{1}{2}} < S(n) < c_1 n^2 - c_3(\varepsilon) n^{(1-\varepsilon)/8}$$

for any $\varepsilon > 0$ (the left-hand inequality is due to R. Alexander).

References

1. G. D. Chakerian and M. S. Klamkin, Minimum triangles inscribed in a convex curve, *Math. Mag.*, 46 (1973).
2. William Lowell Putnam Mathematical Competition, this MONTHLY 76 (1969) 909–915.
3. M. S. Klamkin and D. J. Newman, An inequality for the sum of unit vectors, *Univ. Beo. Publ. Elek. Fac., Ser. Mat. i. Fiz.*, no. 338–352 (1971) 47–48.
4. C. Davis, Theory of positive linear dependence, *Amer. J. Math.*, 76 (1954) 733–746.
5. S. S. Chern, Curves and surfaces in Euclidean space, *Studies in Global Geometry and Analysis*, *Studies in Math.*, Vol. 4, MAA, 1967.
6. H. T. Croft, A net to hold a sphere, *J. London Math. Soc.*, 39 (1964) 1–4.
7. D. S. Mitrinović, *Analytic Inequalities*, Springer-Verlag, Berlin, 1970.
8. G. D. Chakerian and M. S. Klamkin, Minimal covers for closed curves, *Math. Mag.*, 46 (1973) 55–61.
9. J. E. Wetzel, Covering balls for curves of constant length, *L'Enseignement Math.*, 17 (1971) 275–277.
10. D. Laugwitz, Konvexe Mittelpunktsbereiche und normierte Räume, *Math. Z.*, 61 (1954) 235–244.
11. L. Fejes Tóth, *Regular Figures*, Macmillan, New York, 1964.

THE WILLIAM LOWELL PUTNAM MATHEMATICAL COMPETITION

J. H. MCKAY, Oakland University

The following results of the thirty-third William Lowell Putnam Mathematical Competition held on December 2, 1972 have been determined in accordance with the regulations governing the Competition. This competition is supported by the William Lowell Putnam Intercollegiate Memorial Fund left by Mrs. Putnam in memory of her husband and is held under the auspices of the Mathematical Association of America.

The first prize, five hundred dollars, is awarded to the Department of Mathematics of the **California Institute of Technology**, Pasadena, California. The members of the team were Bruce Reznick, Arthur Rubin, and Michael Yoder; to each of these a prize of one hundred dollars is awarded.

The second prize, four hundred dollars, is awarded to the Department of Mathematics of **Oberlin College**, Oberlin, Ohio. The members of the team were Craig Huneke, James Paget, and Craig Seeley; to each of these a prize of seventy-five dollars is awarded.

The third prize, three hundred dollars, is awarded to the Department of Mathematics of **Harvard University**, Cambridge, Massachusetts. The members of the team were David Harbater, David Jerison, and Seth Breidbart; to each of these a prize of fifty dollars is awarded.

The fourth prize, two hundred dollars, is awarded to the Department of Mathe-

matics of **Swarthmore College**, Swarthmore, Pennsylvania. The members of the team were David Hough, David Shucker, and Kin On Tam; to each of these a prize of fifty dollars is awarded.

The fifth prize, one hundred dollars, is awarded to the Department of Mathematics of the **Massachusetts Institute of Technology**, Cambridge, Massachusetts. The members of the team were David Christie, Joseph Mirzoeff, and Scott Brown; to each of these a prize of fifty dollars is awarded.

The six persons ranking highest in the examination, named in alphabetical order, are **Ira Gessel**, Harvard University; **Dean Hickerson**, University of California at Davis; **Arthur Rothstein**, Reed College; **Arthur Rubin**, California Institute of Technology; **David Vogan**, University of Chicago; and **Michael Yoder**, California Institute of Technology. Each of these has been designated as a Putnam Fellow by the Mathematical Association of America and is awarded a prize of two hundred and fifty dollars.

The next four highest ranking individuals, named in alphabetical order, are **Seth Breidbart**, Harvard University; **Paul Lemke**, Rensselaer Polytechnic Institute; **Bruce Reznick**, California Institute of Technology; and **James Shearer**, California Institute of Technology. To each of these a prize of one hundred dollars is awarded.

The following teams, named in alphabetical order, won honorable mention: *Princeton University*, the members of the team were Angelos Tsirimokos, Ray White, and Loring Tu; *Pomona College*, the members of the team were Charles Grinstead, Richard Poppen, and Jerrold Griggs; *Purdue University*, the members of the team were Paul Garrett, Glenn Davis, and Paul Chew; *Rice University*, the members of the team were James Alexander, Gerald Georges, and Edwin Johnson; *University of Toronto*, the members of the team were William Franklin, Peter Debuda, and Robert Anderson.

Honorable mention is given to the following thirty-one individuals, named in alphabetical order: Franklin Adams, *University of Chicago*; Kent Bailey, *Oberlin College*; Scott Brown, *Massachusetts Institute of Technology*; Martin Burger, *Polytechnic Institute of Brooklyn*; David Christie, *Massachusetts Institute of Technology*; Glenn Davis, *Purdue University*; David Dummitt, *California Institute of Technology*; Paul Farmwald, *Purdue University*; Robert Fisher, *California Institute of Technology*; Joseph Grcar, *University of Minnesota*; Alan Grenadir, *Harvard University*; Jerrold Griggs, *Pomona College*; David Hale, *Rensselaer Polytechnic Institute*; David Hough, *Swarthmore College*; Craig Huneke, *Oberlin College*; Glenn Iba, *Massachusetts Institute of Technology*; Paul Ilacqua, *University of Santa Clara*; Thomas Kucera, *University of Manitoba*; Mark Latham, *University of British Columbia*; David Levner, *Cornell University*; Charles Meeker, *Michigan State University*; Brian Mortimer, *Carleton University*; Peter Olver, *Brown University*; James Paget, *Oberlin College*; Robert Rumely, *Grinnell College*; Michael Somos, *Case Western Reserve University*; David Spear, *City College of*

New York; David Ullrich, *University of Wisconsin*; Robert Weissler, *Yale University*; Ray White, *Princeton University*; Chris Wright, *Brown University*.

The other individuals who were ranked in the top one hundred, arranged by college, are: Maria Klawe, *University of Alberta*; John L. Spouge, *University of British Columbia*; Kim Schroeder, *Bucknell University*; Maury Bramson, Howard Landman, and Ross Millikan, *University of California at Berkeley*; Lawrence Gray, *University of California at San Diego*; Richard Niles, *University of California at Santa Cruz*; Thomas Howell, *California Institute of Technology*; Stewart Strait, *California State University at San Diego*; Michael Rennie, *California State University at San Jose*; Robert Bundy and Steven Kalikow, *Case Western Reserve University*; Thomas Branson and James McClure, *University of Chicago*; Charles Levermore, *Clarkson College*; Meir Shinnar and Jacob Sturm, *Columbia University*; Daniel Fisher and Jeffrey Hoffstein, *Cornell University*; David Garlock, David Jerison, Bruce Leverett, David Mostow, and Lyle Ramshaw, *Harvard University*; John Boyd, *Indiana University*; Robert Beggs, *Kent State University*; John Bate, *University of Manitoba*; Dale Johannesen, *University of Maryland*; Leo Katzenstein and James Marlin, *Massachusetts Institute of Technology*; Dennis Stowe and David Catlin, *University of Michigan*; Steve Eicker and John Reiser, *Michigan State University*; Robert Bradley and Jean-Louis Richer, *Université de Montréal*; Albert Wigchert, *University of Nevada at Reno*; Eric Lofgren, *New College*; John Gilbert, *University of New Mexico*; Brian Hulse, *New York University*; Eli Isaacson, *New York University, Washington Square*; Craig Seeley, *Oberlin College*; Julius Collins, *Polytechnic Institute of Brooklyn*; Angelos Tsirimokos and Stephen Weber, *Princeton University*; Daniel Bump, *Reed College*; James Alexander, Edwin Johnson, and Joseph Robinett, *Rice University*; John Hyde and John Mellby, *St. Olaf College*; Jerome Eastham, *Southwestern at Memphis*; David Shucker, *Swarthmore College*; Robert Anderson and William Franklin, *University of Toronto*; Lanh Dinh Dang, *University of Utah*; William Parke, *University of Washington*; Jan Verster, *University of Waterloo*; Matthew Ginsberg, *Wesleyan University*; Robert Marheine and Robert Mortenson, *University of Wisconsin at Madison*; David Morandi, *University of Wyoming*.

One thousand six hundred and eighty one students from three hundred and twenty two colleges and universities in the United States and Canada participated in the examination on December 2, 1972.

The Questions Committee, consisting of Murray Klamkin (chairman), Nathan S. Mendelsohn, and Donald J. Newman, prepared the problems (listed below) for the competition.

PROBLEMS. PART A

- A-1. Show that there are no four consecutive binomial coefficients $\binom{n}{r}, \binom{n}{r+1}, \binom{n}{r+2}, \binom{n}{r+3}$ (n, r integers > 0 and $r+3 \leq n$) which are in arithmetic progression.

- A-2. Let S be a set and let $*$ be a binary operation on S satisfying the laws

$$x * (x * y) = y \quad \text{for all } x, y \text{ in } S,$$

$$(y * x) * x = y \quad \text{for all } x, y \text{ in } S.$$

Show that $*$ is commutative but not necessarily associative.

- A-3. If for a sequence x_1, x_2, x_3, \dots , $\lim_{n \rightarrow \infty} (x_1 + x_2 + \dots + x_n)/n$ exists, call this limit the C -limit of the sequence. A function $f(x)$ from $[0, 1]$ to the reals is called a supercontinuous function on the interval $[0, 1]$ if the C -limit exists for the sequence $f(x_1), f(x_2), f(x_3), \dots$ whenever the C -limit exists for the sequence x_1, x_2, x_3, \dots . Find all supercontinuous functions on $[0, 1]$.
- A-4. Of all ellipses inscribed in a square, show that the circle has the maximum perimeter.
- A-5. Show that if n is an integer greater than 1, then n does not divide $2^n - 1$.
- A-6. Let $f(x)$ be an integrable function in $0 \leq x \leq 1$ and suppose $\int_0^1 f(x)dx = 0$, $\int_0^1 x f(x)dx = 0, \dots$, $\int_0^1 x^{n-1} f(x)dx = 0$ and $\int_0^1 x^n f(x)dx = 1$. Show that $|f(x)| \geq 2^n(n+1)$ in a set of positive measure.

PART B

- B-1. Show that the power series representation for the series $\sum_{n=0}^{\infty} (x^n(x-1)^{2n})/n!$ cannot have three consecutive zero coefficients.
- B-2. A particle moving on a straight line starts from rest and attains a velocity v_0 after traversing a distance s_0 . If the motion is such that the acceleration was never increasing, find the maximum time for the traverse.
- B-3. Let A and B be two elements in a group such that $ABA = BA^2B$, $A^3 = 1$ and $B^{2n-1} = 1$ for some positive integer n . Prove $B = 1$.
- B-4. Let n be an integer greater than 1. Show that there exists a polynomial $P(x, y, z)$ with integral coefficients such that $x \equiv P(x^n, x^{n+1}, x + x^{n+2})$.
- B-5. If the opposite angles of a skew (non-planar) quadrilateral are equal in pairs, prove that the opposite sides are equal in pairs.
- B-6. Let $n_1 < n_2 < n_3 < \dots < n_k$ be a set of positive integers. Prove that the polynomial $1 + z^{n_1} + z^{n_2} + \dots + z^{n_k}$ has no roots inside the circle $|z| < (\sqrt{5} - 1)/2$.

SOLUTIONS. PART A

The number in parentheses, immediately following the problem number, is the number of participants who received a score of 8, 9 or 10 (10 is maximum possible) on the problem. In the case of A-1 A-2, B-1, and B-2, this applies to all 1681 participants. For the other problems, the count applies only to the 957 qualifiers.

A-1 (129). For a given n and r , in order for the first three binomial coefficients to be in arithmetic progression, we must have

$$(1) \qquad 2\binom{n}{r+1} = \binom{n}{r} + \binom{n}{r+2}$$

or equivalently

$$(2) \qquad 2 = \frac{r+1}{n-r} + \frac{n-r-1}{r+2}.$$

The condition that the last three given binomial coefficients are in arithmetic progression is found from (1) by replacing r by $r+1$. Consequently both r and $r+1$ must satisfy equation (2) if all four terms are in arithmetic progression.

Note that the two terms in equation (2) are interchanged if r is replaced by $n-r-2$. Thus the quadratic equation (2) has roots

$$r, r+1; n-r-3, n-r-2.$$

Since (2) can have only two roots, $r = n-r-3$ and $n = 2r+3$. The four binomial coefficients must be

$$\binom{2r+3}{r}, \binom{2r+3}{r+1}, \binom{2r+3}{r+2}, \binom{2r+3}{r+3}$$

which are the four middle terms. They cannot be in arithmetic progression since binomial coefficients increase to the middle term(s) and then decrease.

A-2 (167). Label the given laws (1) and (2), respectively.

I. We first show that

$$(3) \qquad (x * y) * x = y.$$

This follows from $(x * y) * x = (x * y) [(x * y) * y] = y$. (First apply (2) with x and y interchanged; then apply (1) with x replaced by $x * y$.)

We now obtain

$$(4) \qquad y * x = [(x * y) * x] * x = x * y.$$

(First apply (3); then apply (2) with y replaced by $x * y$.) This proves that $*$ is commutative.

II. Let S be the set of all integers. Define $x * y = -x - y$. Then

$$(5) \qquad x * (y * z) = -x + y + z; (x * y) * z = x + y - z.$$

It follows from (5) that, in the first place, (1) and (2) hold and, secondly, $*$ fails to be associative: simply choose $x \neq z$ in (5).

Alternate Solution, Part I (suggested by Martin Davis):

Write the equation $x * y = z$ as $P(x, y, z)$. Then law (1) may be written "If $x * y = z$

then $x * z = y$ '' or

$$(6) \quad P(x, y, z) \text{ implies } P(x, y, z).$$

Similarly, the law (2) may be written

$$(7) \quad P(y, x, z) \text{ implies } P(z, x, y).$$

These two implications, (6) and (7), show that the permutations (23) and (13) on the location of the variables in $P(x, y, z)$ are permitted. Since (13), (23) generate the symmetric group S_3 , we find (12) is also permitted.

Thus, $P(x, y, z)$ implies $P(y, x, z)$ or $x * y = z$ implies $y * x = z$, which means $x * y = y * x$.

A-3 (6). A function is "supercontinuous" if and only if it is affine, $f(x) = Ax + B$. The sufficiency is trivial (and was worth 1 point in the grading). For the necessity: First we note that it is *not* assumed that $f(C\text{-limit}) = C\text{-limit}(f)$ (otherwise the solution could be materially simplified). The essential steps are to show, that if f is supercontinuous, then (1) f is continuous, and (2) $f((a+b)/2) = (f(a) + f(b))/2$ for all a, b . These two statements imply that f is affine. The proofs of (1) and (2) are similar; we give (2) (which is the harder). Set $c = (a+b)/2$, and suppose $f(c) \neq (f(a) + f(b))/2$. Imagine any sequence of integers N_i which "grows very rapidly"; say let N_{i+1} exceed $2^i N^i$. Then construct a sequence of points $\{x_n\}$ as follows: Break the sequence into blocks, alternating between

$$\begin{aligned} \{x_n\} &= a, b, a, b, a, b, \dots \\ \{x_n\} &= c, c, c, c, c, c, \dots, \end{aligned}$$

the ab pattern holding for $N_{2i-1} \leq n < N_{2i}$, and the c pattern holding for $N_{2i} \leq n < N_{2i+1}$. Then $\{x_n\}$ has the C -limit c , but the averages of $\{f(x_n)\}$ oscillate (because the lengths of the blocks $N_i \leq n < N_{i+1}$ increase very fast, and $f(c) \neq$ the average of $f(a)$ and $f(b)$). Thus the C -limit of $\{f(x_n)\}$ does not exist, a contradiction.

Comments: Many interesting classes of functions were proposed as the "answers" to this question. The most common choices were the class of all bounded functions and the continuous functions. (The correct choice ranked third in frequency.) Riemann and Lebesgue integrable functions were also mentioned.

Professor David Cohoon has suggested the following problem: Is there any topology on the real line in terms of which the class of continuous functions coincides with the "supercontinuous" functions? (Here the functions are from R to R , and the same topology is to be put on R as image space and domain.)

A-4 (1). Let the square of sidelength $2R$ have the vertices $(\pm R\sqrt{2}, 0)$ and $(0, \pm R\sqrt{2})$. The ellipse

$$(1) \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

with $0 \leq b \leq a \leq R\sqrt{2}$ has the line $x + y = R\sqrt{2}$ as a tangent if and only if the quadratic equation $x^2/a^2 + (R\sqrt{2} - x)^2/b^2 = 1$ has a double root. It can be verified that its discriminant vanishes if and only if $a^2 + b^2 = 2R^2$. As a varies from R to $R\sqrt{2}$ and b varies from R to 0, the curve (1) varies from the circle of radius R through all the non-circular ellipses inscribed in the square to the degenerate "flat" ellipse lying on the x -axis.

Let $4L$ denote the length of the ellipse $x = a \cos t$, $y = b \sin t$, $0 \leq t \leq 2\pi$. Then

$$\begin{aligned} L &= \int_0^{\pi/2} [a^2 \sin^2 t + b^2 \cos^2 t]^{\frac{1}{2}} dt = \int_0^{\pi/2} [\tfrac{1}{2}a^2(1 - \cos 2t) + \tfrac{1}{2}b^2(1 + \cos 2t)]^{\frac{1}{2}} dt \\ &= \int_0^{\pi/2} [R^2 - \tfrac{1}{2}c^2 \cos 2t]^{\frac{1}{2}} dt, \end{aligned}$$

where $c^2 = a^2 - b^2$. The last integral we split into one from 0 to $\pi/4$ and one from $\pi/4$ to $\pi/2$, and in the latter we substitute $t = \pi/2 - t'$, obtaining

$$(2) \quad L = \int_0^{\pi/4} \{[R^2 - \tfrac{1}{2}c^2 \cos 2t]^{\frac{1}{2}} + [R^2 + \tfrac{1}{2}c^2 \cos 2t]^{\frac{1}{2}}\} dt.$$

Note that $\cos 2t > 0$ for $0 \leq t < \pi/4$.

Now the function $f(u) = (p - u)^{\frac{1}{2}} + (p + u)^{\frac{1}{2}}$ decreases in the interval $0 \leq u \leq p$, because $2f'(u) = -(p - u)^{-\frac{1}{2}} + (p + u)^{-\frac{1}{2}} < 0$ for $0 < u < p$. Thus the integral in (2) as a function of c has its largest value when $c = 0$, that is, for the inscribed circle.

To show that an ellipse inscribed in the square must have its axes along the diagonals of the square, we choose the square as having sides $u = \pm R$ and $v = \pm R$ and the ellipse as having the equation

$$Au^2 + Buv + Cv^2 + Du + Ev + F = 0,$$

where

$$(1) \quad 4AC - B^2 > 0.$$

Taking the "highest," "lowest," "rightest," and "leftest" points on the ellipse, we see that all four sides of the square must be tangents to the ellipse.

The line $u = R$ is a tangent if and only if the equation $Cv^2 + (BR + E)v + (AR^2 + DR + F) = 0$ has a double root or

$$(2) \quad (BR + E)^2 - 4C(AR^2 + DR + F) = 0.$$

The corresponding conditions for $u = -R$, $v = R$ and $v = -R$ are

$$(3) \quad (-BR + E)^2 - 4C(AR^2 - DR + F) = 0,$$

$$(4) \quad (BR + D)^2 - 4A(CR^2 + ER + F) = 0,$$

$$(5) \quad (-BR + D)^2 - 4A(CR^2 - ER + F) = 0,$$

respectively. Subtract (2) from (3) and divide by $4R$; this gives

$$(6) \quad 2CD - BE = 0.$$

Similarly, from (4) and (5),

$$(7) \quad -BD + 2AE = 0.$$

By (6), (7) and (1), $D = E = 0$. Therefore (2) and (4) become

$$B^2R^2 - 4ACR^2 - 4CF = 0, \quad B^2R^2 - 4ACR^2 - 4AF = 0,$$

respectively. Since $F \neq 0$, we have $A = C$; this means that the ellipse has its axes along the lines $u \pm v = 0$.

A-5 (22). Assume that n divides $2^n - 1$ for some $n > 1$. Since $2^n - 1$ is odd, n is odd. Let p be the smallest prime factor of n . By Euler's Theorem, $2^{\phi(p)} \equiv 1 \pmod{p}$, because p is odd. If λ is the smallest positive integer such that $2^\lambda \equiv 1 \pmod{p}$ then λ divides $\phi(p) = p - 1$. Consequently λ has a smaller prime divisor than p . But $2^n \equiv 1 \pmod{p}$ and so λ also divides n . This means that n has a smaller prime divisor than p . Contradiction.

A-6 (9). The conditions imply $\int_0^1 (x - \frac{1}{2})^n f(x) dx = 1$. Suppose $|f(x)| < 2^n(n+1)$ except for a set of measure 0.

Then $1 = \int_0^1 (x - \frac{1}{2})^n f(x) dx < 2^n(n+1) \int_0^1 |x - \frac{1}{2}|^n dx = 1$, a contradiction.

SOLUTIONS. PART B

B-1 (45). For the proposed solution the problem could have been stated in the more general form: The series expansion about any point for $\exp(P(x))$, if $P(x)$ is a cubic polynomial, will not have three consecutive zero coefficients.

If $f(x) = \exp(P(x))$, where $P(x)$ is a cubic polynomial, then $f' = f \cdot P'$ and $f'' = f' \cdot P' + f \cdot P''$. In general for $k \geq 2$,

$$(1) \quad f^{(k+1)} = f^{(k)} \cdot P' + \binom{k}{1} f^{(k-1)} \cdot P'' + \binom{k}{2} f^{(k-2)} \cdot P'''.$$

It follows from (1): if, at some (real or complex) point x_0 , $f^{(k-2)}(x_0) = f^{(k-1)}(x_0) = f^{(k)}(x_0) = 0$, then also $f^{(k+1)}(x_0) = 0$. By the same argument, $f^{(\mu)}(x_0) = 0$ for $\mu = k+2, k+3, \dots$; so that $f(x)$ would reduce to a polynomial. This is evidently impossible.

Alternate Solution: In the given form of the problem it can be shown that no coefficient of x^k is zero. The product $x^n(1-x)^{2n}$ has a non-zero coefficient for x^k if $0 \leq k-n \leq 2n$ or, equivalently, $k/3 \leq n \leq k$. This coefficient is the integer

$$(-1)^{k-n} \binom{2n}{n-k},$$

which we denote by $a(n, k)$. The coefficient of x^k in the given series is

$$C_k = \sum_{n=\lfloor k/3 \rfloor + 1}^k \frac{a(n, k)}{n!}.$$

Multiplying through this summation by $(k-1)!$ will convert each term, except the last term, to an integer. The last term becomes $1/k$. Since $(k-1)!$ times C_k is not an integer for $k > 1$ and $C_1 = C_0 = 1$, there are no zero coefficients in the expansion of the given series in powers of x .

B-2 (152). We take v_0 as positive (see Comment) and consider the graph of v as a function of t (see Figure 1). From the given data we know that the curve starts at the origin and is concave downward since the acceleration $a = dv/dt$ does not increase.

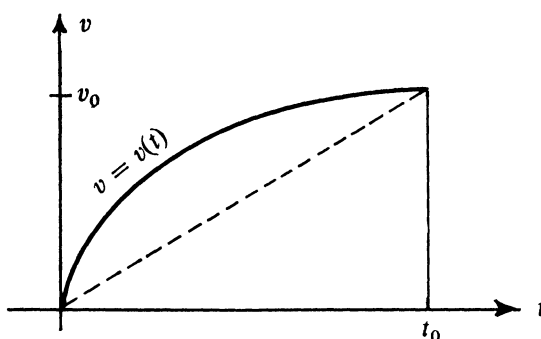


FIG. 1

Let t_0 be the time of the traverse. Then $v(t_0) = v_0$. The distance s_0 is represented by the area bounded by the curve $v = v(t)$, the t -axis, and the line $t = t_0$. The area of the right triangle with vertices at $(0, 0)$, $(t_0, 0)$ and (t_0, v_0) has area less than or equal to s . Thus $\frac{1}{2} v_0 t_0 \leq s_0$ or

$$t_0 \leq \frac{2s_0}{v_0}.$$

Equality is possible and gives the maximum value of t_0 (for given s_0 and v_0) when the graph of $v(t)$ is the straight line $v(t) = (v_0/t_0)t = (v_0^2/2s_0)t$.

Comment: If v_0 is zero or negative, there is no maximum time t_0 for the traverse. In the case $v_0 = 0$ the equation of motion

$$S = s_0[3(t/t_0)^2 - 3(t/t_0)^3], \quad 0 \leq t \leq t_0$$

satisfies the conditions of the problem for any $t_0 > 0$.

B-3 (123). From $ABA = BA^2B = BA^{-1}B$, we have

$$AB^2 = ABA \cdot A^{-1}B = BA^{-1}BA^{-1}B = BA^{-1} \cdot ABA = B^2A.$$

By induction, $AB^{2r} = B^{2r}A$ so that $AB = AB^{2n} = B^{2n}A = BA$. Since A and B commute, $ABA = BA^2B$ implies $A^2B = A^2B^2$, or $B = B^2$, or $B = 1$.

Alternate Solution: It can be shown that A and B commute by expressing each as powers of the same group element. Because $A^3 = 1$ it is tempting to multiply $ABA = BA^2B$ on the right by A^2 and then on the left by BA^2 to get $B^2 = (BA^2)^3$. Set $X = BA^2$ and use $B^{2n} = B$ to obtain

$$(1) \quad B = X^{3n}.$$

From $X = BA^2$, we get $XA = B$, $A = X^{-1}B$, or

$$(2) \quad A = X^{3n-1}.$$

The conclusion that $B = 1$ is as before.

B-4 (99). Let $x = t^n$, $y = t^{n+1}$, $z = t + t^{n+2}$. We construct a polynomial $P(x, y, z)$ with integral coefficients such that $P(x, y, z) = t$. We have

$$\begin{aligned} z &= t + t^{n+2}, \\ zy &= t^{n+2} + t^{2n+3}, \\ zy^2 &= t^{2n+3} + t^{3n+4}, \\ &\dots \dots \dots \\ z^{n-2} &= t^{n^2-n-1} + t^{n^2}. \end{aligned}$$

Multiply the above equations alternately by $+1$ and -1 and add:

$$z[1 - y + y^2 - \dots + (-1)^{n-2}y^{n-2}] = t + (-1)^{n-2}t^{n^2} = t + (-1)^n x^n.$$

Hence, if we define

$$P(x, y, x) = z \left[\sum_{i=0}^{n-2} (-1)^i y^i \right] + (-1)^{n-1} x^n,$$

Then $P(t^n, t^{n+1}, t + t^{n+2}) = t$.

B-5 (0). For the skew quadrilateral $ABCD$, let $AB = a$, $BC = b$, $CD = c$, $DA = d$, $AC = x$, $BD = y$. None of these lengths can be zero. By the law of cosines:

$$\frac{a^2 + b^2 - x^2}{ab} = \frac{c^2 + d^2 - x^2}{cd}$$

or $(ab - cd)x^2 = (bc - ad)(ac - bd)$. Similarly, $(ad - bc)y^2 = (cd - ab)(ac - bd)$.

CASE 1: $ab - cd = 0$.

Then, $ad - bc = 0$ and $a = c$, $b = d$.

CASE 2: $ab - cd \neq 0$.

Then, $bc - ad \neq 0$, $ac - bd \neq 0$ and $x^2y^2 = (ac - bd)^2$. Consequently,

$$ac = xy + bd \text{ or } bd = ac + xy.$$

By Ptolemy's Theorem (in space), $ABCD$ must be concyclic which violates the skew condition.

Alternate Solution: If $AC = BC$ and $AD = BD$, the conclusion that $AC = BD$ and $BC = AD$ is obvious (see Figure 2) so assume $AC \neq BC$. With this assumption we first show $BD = AC$. If $BD \neq AC$ there exists a unique point D^* in the plane of $\triangle ADB$ with $BD^* = AC$, $AD^* = CB$. $\angle AD^*B = \angle ACB = \theta$. From \triangle 's ADE and BD^*E it follows that $\angle DAE = \angle D^*BE$.

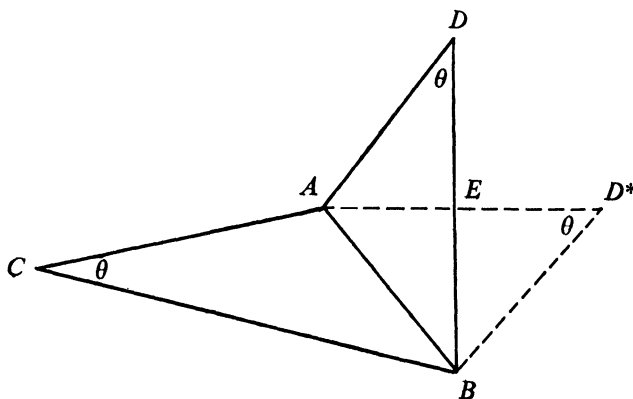


FIG. 2

From the congruent \triangle 's CD^*A and CD^*B it follows that $\angle CAD^* = \angle CBD^*$. These angle equalities prove that the trihedral angles $A - CDD^*$ and $B - CDD^*$ are congruent. Hence the angle which \vec{CA} makes with the plane ADD^* is equal to the angle \vec{CB} makes with the plane BDD^* (which is the plane ADD^*). If H is the foot of the altitude from C to this plane the \triangle 's CHA and CHB are congruent right triangles. That is $AC = CB$. This is a contradiction and $BD = AC$.

Interchanging the roles of B and A in the above shows that $AD = BC$.

B-6 (1). Let $P(z)$ denote the given polynomial. The power series expansion of $1/(1-z) - 2P(z)$ has coefficients ± 1 with leading coefficient -1 . Hence,

$$(1) \quad \left| 1 + \frac{1}{1-z} - 2P(z) \right| \leq |z| + |z|^2 + \cdots = \frac{|z|}{1-|z|}.$$

Also,

$$\begin{aligned} |2P(z)| &\geq \left| 1 + \frac{1}{1-z} \right| - \left| 1 + \frac{1}{1-z} - 2P(z) \right| \\ &\geq 1 + \frac{1}{1+|z|} - \frac{|z|}{1-|z|} = 2 \frac{1-|z| - |z|^2}{1-|z|^2}. \end{aligned}$$

The latter term is positive for $|z| < (\sqrt{5} - 1)/2$.

Acknowledgments

The Director acknowledges, with appreciation, the assistance of Fritz Herzog, the Questions Committee, and the graders, especially L. M. Kelly, M. Hausner and J. I. Richards, in preparing the above solutions. The graders for the competition were: J. C. Chipman, C. V. Coffman, J. W. Dettman, R. A. DeVore, R. M. Dudley, W. R. Emerson, D. J. Eustice, R. A. Fontenot, J. Froemke, R. A. Gambill, M. Hausner, A. P. Hillman, L. M. Kelly, B. B. Lieberman, D. G. Malm, E. A. Nordhaus, R. Pollack, J. I. Richards, D. Rosen, P. J. Sally, M. E. Shanks, H. A. Smith, K. E. Westerbeck, E. T. Wong.

SIMPLE GROUPS*

(Sung to the tune of "Sweet Betsy from Pike")

What are the orders of all simple groups? I speak of the honest ones, not of the loops. It seems that old Burnside their orders has guessed Except for the cyclic ones, even the rest. Groups made up with permutes will produce some more: For A_n is simple, if n exceeds 4. Then, there was Sir Matthew who came into view Exhibiting groups of an order quite new. Still others have come on to study this thing. Of Artin and Chevalley now we shall sing. With matrices finite they made quite a list The question is: Could there be others they've missed? Suzuki and Ree then maintained it's the case That these methods had not reached the end of the chase. They wrote down some matrices, just four by four, That made up a simple group. Why not make more? And then came the opus of Thompson and Feit Which shed on the problem remarkable light.	A group, when the order won't factor by two Is cyclic or solvable. That's what is true. Suzuki and Ree had caused eyebrows to raise, But the theoreticians they just couldn't faze. Their groups were not new: if you added a twist, You could get them from old ones with a flick of the wrist. Still, some hardy souls felt a thorn in their side. For the five groups of Mathieu all reason defied; Not A_n , not twisted, and not Chevalley, They called them sporadic and filed them away. Are Mathieu groups creatures of heaven or hell? Zvonimir Janko determined to tell. He found out that nobody wanted to know: The masters had missed 1 7 5 5 6 0. The floodgates were opened! New groups were the rage! (And twelve or more sprouted, to greet the new age.) By Janko and Conway and Fischer and Held McLaughlin, Suzuki, and Higman, and Sims. No doubt you noted the last lines don't rhyme. Well, that is, quite simply, a sign of the time. There's chaos, not order, among simple groups; And maybe we'd better go back to the loops.
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

* Found scrawled on a library table in Eckhart Library at the U. of Chicago; author unknown, or in hiding. (See W. E. Mientka, Professor Leo Moser — Reflections of a Visit, *American Mathematical Monthly* 79 (1972), 609–614.)

MATHEMATICAL NOTES

EDITED BY ROBERT GILMER

Material for this Department should be sent to David Roselle, Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803.

ON A PROBLEM CONCERNING EULER'S PHI-FUNCTION

HAROLD DONNELLY, Berkeley, California

Let $\phi(N)$ denote Euler's phi-function, thus $\phi(N)$ is the number of positive integers less than N which are relatively prime to N . Our attention is centered on whether there exists a number M such that $\phi(N) = M$ has exactly one solution [cf. 2, p. 37, prob. 15].

While this problem is unsolved at present we obtain several results, including a lower bound on the size the integer N must have.

The well-known formula $\phi(\pi\rho_i^{\alpha_i}) = \pi\rho_i^{\alpha_i-1}(\rho_i-1)$ will be applied throughout. The table of prime numbers given by Baker and Gruenberger [1] was used to verify that certain numbers appearing in the proofs of Theorems 3 and 4 are primes as asserted.

THEOREM 1. *If N is the unique solution to $\phi(N) = M$ for some M , then $2^23^27^243^2$ divides N .*

Proof. If $2 \nmid N$, then $\phi(2N) = \phi(N)$ and if $2 \mid N$ but $2^2 \nmid N$, then $\phi(2^{-1}N) = \phi(N)$. Thus $2^2 \mid N$. If $3 \nmid N$, then $\phi(2^{-1}3N) = \phi(N)$; if $3 \mid N$ and $3^2 \nmid N$, then $\phi(2^13^{-1}N) = \phi(N)$. Thus we have so far that $2^23^2 \mid N$. Next, if $7 \nmid N$, then $\phi(2^{-1}3^{-1}7N) = \phi(N)$; if $7 \mid N$ and $7^2 \nmid N$, then $\phi(2^13^17^{-1}N) = \phi(N)$, and therefore $2^23^27^2 \mid N$. Finally if $43 \nmid N$, then $\phi(2^{-1}3^{-1}7^{-1}43N) = \phi(N)$; if $43 \mid N$ and $43^2 \nmid N$, then $\phi(2^13^17^143^{-1}N) = \phi(N)$, and so $2^23^27^243^2 \mid N$.

NOTE. The proof clearly rests on the fact that each prime divisor is the product of previous prime divisors plus one, i.e., $3 = 2 + 1$, $7 = 2 \times 3 + 1$, $43 = 2 \times 3 \times 7 + 1$. This cannot be continued since the remaining products of known prime divisors plus one are $15 = 2 \times 7 + 1$, $87 = 2 \times 43 + 1$, $259 = 2 \times 3 \times 43 + 1$, $603 = 2 \times 7 \times 43 + 1$, $1807 = 2 \times 3 \times 7 \times 43 + 1$ and these are all composite.

THEOREM 2. *If K is the smallest N such that for some M $\phi(N) = M$ has exactly one solution, then 2^3 does not divide K .*

Proof. By Theorem 1, $2^2 \mid K$. Since K is the smallest number giving exactly one solution we know there is a $J \neq 2^{-1}K$ such that $\phi(J) = \phi(2^{-1}K)$. If $2 \mid J$, then $\phi(2J) = \phi(K)$ implying $J = 2^{-1}K$ since K is a unique solution; therefore $2 \nmid J$. Thus $\phi(4J) = 2\phi(J) = 2\phi(2^{-1}K) = \phi(K)$ and we see that $J = 4^{-1}K$. But if $2^3 \mid K$,

then $\phi(4^{-1}K) = 4^{-1}\phi(K)$; but $\phi(2^{-1}K) = 2^{-1}\phi(K)$ and these are not equal, therefore $2^3 \nmid K$.

REMARK. By Theorem 1, it follows that the second power is the exact power of 2 dividing K .

We next consider the two possibilities for 3; i.e., $3^3 \mid K$ and $3^3 \nmid K$.

THEOREM 3. *If K is the smallest number N such that for some M $\phi(N) = M$ has exactly one solution and it should occur that 3^3 divides K , then $(3)(2)(3)(7)(19)(43)(127)(2287)(4903)(5419)(14,479)(98,299)(101,347)(304,039)(617,767)(688,087)(4,324,363)(26,563,939)(78,456,283)(86,714,839)]^2$ divides K .*

Proof. From Theorem 1 and our hypotheses, $2^2 3^3 7^2 43^2 \mid K$.

If $19 \nmid K$, then $\phi(18^{-1}19K) = \phi(K)$. If $19 \mid K$ and $19^2 \nmid K$, then $\phi(18^1 19^{-1}K) = \phi(K)$, and so $19^2 \mid K$. Similarly if $127 \nmid K$, then $\phi(126^{-1}127K) = \phi(K)$. If $127 \mid K$ and $127^2 \nmid K$, then $\phi(126^1 127^{-1}K) = \phi(K)$; if $5419 \nmid K$, then $\phi(5418^{-1}5419K) = \phi(K)$; if $5419 \mid K$ but $5419^2 \nmid K$, then $\phi(5418^1 5419^{-1}K) = \phi(K)$. Note that this process works because $19 = 2 \times 3^2 + 1$, $127 = 2 \times 3^2 \times 7 + 1$; $5419 = 2 \times 3 \times 7 \times 43 + 1$ and because 19, 127, 5419 are prime (cf., [1]).

We thus have so far that $2^2 3^3 7^2 19^2 43^2 127^2 5419^2 \mid K$.

We next note that

$$2287 = 2 \times 3^2 \times 127 + 1; 4903 = 2 \times 3 \times 19 \times 43 + 1;$$

$$14,479 = 2 \times 3 \times 19 \times 127 + 1; 98,299 = 2 \times 3^2 \times 43 \times 127 + 1;$$

$$101,347 = 2 \times 3 \times 7 \times 19 \times 127 + 1; 304,039 = 2 \times 3^2 \times 7 \times 19 \times 127 + 1;$$

$$617,767 = 2 \times 3 \times 19 \times 5419 + 1; 688,087 = 2 \times 3^2 \times 7 \times 43 \times 127 + 1;$$

$$4,324,363 = 2 \times 3 \times 7 \times 19 \times 5419 + 1; 26,563,939 = 2 \times 3 \times 19 \times 43 \times 5419 + 1;$$

$$78,456,383 = 2 \times 3 \times 19 \times 127 \times 5419 + 1; 86,714,839 = 2 \times 3^2 \times 7 \times 127 \times 5419 + 1;$$

and that the numbers appearing on the left hand side of the equal signs are prime (cf. [1]). The theorem results by arguments entirely similar to those above.

THEOREM 4. *If K is the smallest number N such that for some M $\phi(N) = M$ has exactly one solution and it should occur that $3^3 \nmid K$, then $[(2)(3)(7)(13)(43)(79)(157)(547)(1093)(3613)(6709)(46,957)(303,493)(12,118,003)]^2$ divides K .*

Proof. $2^2 3^2 7^2 43^2$ divides K by Theorem 1, $2^3 \nmid K$ by Theorem 2, and $3^3 \nmid K$ by assumption.

If $13 \nmid K$, then $\phi(13^1 36^{-1}K) = \phi(K)$ so $13 \mid K$. If $13 \mid K$, $13^2 \nmid K$, then $\phi(12^1 13^{-1}K) = \phi(K)$. So $13^2 \mid K$. If $3613 \nmid K$, then $\phi(7^{-1} 36^{-1} 43^{-1} 3613K) = \phi(K)$. If $3613 \mid K$, $3613^2 \nmid K$, then $\phi(3612^1 3613^{-1}K) = \phi(K)$. So $2^2 3^2 7^2 13^2 43^2 3613^2 \mid K$.

This process works because $13 = 2^2 \times 3 + 1$; $3613 = 2^2 \times 3 \times 7 \times 43 + 1$ and because 13, 3613 are prime (cf. [1]).

Noting that

$$79 = 2 \times 3 \times 13 + 1; 157 = 2^2 \times 3 \times 13 + 1; 547 = 2 \times 3 \times 7 \times 13 + 1;$$

$$1093 = 2^2 \times 3 \times 7 \times 13 + 1; 6709 = 2^2 \times 3 \times 13 \times 43 + 1;$$

$$46,957 = 2^2 \times 3 \times 7 \times 13 \times 43 + 1; 303,493 = 2^2 \times 3 \times 7 \times 3613 + 1;$$

$$12,118,003 = 2 \times 3 \times 13 \times 43 \times 3613 + 1;$$

and that the numbers appearing on the left hand side of the equations are prime (cf. [1]) we obtain the required result.

THEOREM 5. *A lower bound on a number N such that $\phi(N) = M$ has exactly one solution is 10^{77} .*

Proof. Follows from Theorems 3 and 4.

References

1. C. L. Baker, and F. J. Gruenberger, The first Six Million Prime Numbers, Santa Monica, The Rand Corporation, 1957.
2. Ivan Niven and H. S. Zuckerman, An Introduction to the Theory of Numbers, Wiley, New York, 1964.

WHAT IS THE PROBABILITY THAT TWO GROUP ELEMENTS COMMUTE?

W. H. GUSTAFSON, Indiana University

1. Introduction. A student studying both probability and algebra might well ask the question posed in the title. One can solve the problem for finite groups by a straightforward attack as follows:

Let G be a group of finite order n . The probability $\text{Pr}(G)$ that two elements selected at random (with replacement) from G are commutative is $|C|/n^2$, where $C = \{(x, y) \in G \times G \mid xy = yx\}$. In order to count the elements of C , we observe that for each $x \in G$, the number of elements of C of the form (x, y) is $|C_x|$, where C_x is the centralizer of x in G . Hence we have

$$|C| = \sum |C_x|,$$

where the sum extends over all $x \in G$. Now we recall that if x and y are conjugate elements of G , then C_x and C_y are conjugate subgroups. Further, the number of elements in the conjugacy class of x is $[G : C_x]$. Hence, if x_1, \dots, x_k are representatives of the conjugacy classes in G , we have

$$|C| = \sum_{i=1}^k [G : C_{x_i}] \cdot |C_{x_i}| = k \cdot n.$$

Thus $\text{Pr}(G) = k/n$, the number of classes in G divided by the order of G . This technique was used by Erdős and Turán [4].

Now let us observe that $5/8$ is an upper bound for $\text{Pr}(G)$ when G is nonabelian. For the "class equation" tells us that

$$|G| = |Z| + |K_1| + \cdots + |K_t|,$$

where Z is the center of G , and K_1, \dots, K_t are the nontrivial conjugacy classes. We have $|K_i| \geq 2$ for $i = 1, \dots, t$, whence $(|G| - |Z|)/2 \geq t$. Thus $k = t + |Z| \leq (|G| + |Z|)/2$. As G is nonabelian, G/Z is not cyclic (see Scott [7, p. 50]) and hence $|Z| \leq |G|/4$. Thus $k \leq 5/8 \cdot |G|$, so $\text{Pr}(G) \leq 5/8$. The reader may verify that this bound is sharp, by examining the nonabelian groups of order eight.

2. Compact groups. The reader may now wonder whether the above analysis carries over in any sense to infinite groups. Of course, the ratio k/n is no longer meaningful, but we shall see that there is an analogue of the bound $5/8$ for a class of topological groups.

Let G be a compact, Hausdorff topological group. We recall that G has a **left Haar measure**; that is, a Borel measure μ such that $\mu(U) > 0$ for each nonempty open set U of G , and $\mu(x \cdot E) = \mu(E)$ for each Borel set E of G and each $x \in G$. Further, μ is unique once we impose the normalization condition $\mu(G) = 1$. The reader who is not familiar with Haar measure may consult [5, Chapter XI]. On the product space $G \times G$, we impose the product measure $\mu \times \mu$. Again let $C = \{(x, y) \in G \times G \mid xy = yx\}$. We remark that $C = f^{-1}(1)$, where $f: G \times G \rightarrow G$ is the continuous function given by $f(x, y) = xyx^{-1}y^{-1}$. It follows that C is closed, and hence measurable. We view $\mu \times \mu$ as a probability measure; then $\text{Pr}(G) = \mu \times \mu(C)$. Let us now prove our generalization of the last result of Section 1:

THEOREM. *Let G be a compact nonabelian group. Then $\text{Pr}(G) \leq 5/8$.*

Proof. Let $\chi: G \times G \rightarrow \text{reals}$, be the characteristic function of C . Then we have

$$\mu \times \mu(C) = \int_{G \times G} \chi d(\mu \times \mu).$$

By Fubini's theorem [5, p. 148], we have

$$\mu \times \mu(C) = \int_G \int_G \chi(x, y) d\mu(y) d\mu(x).$$

Also, $\int_G \chi(x, y) d\mu(y) = \mu(C_x)$ for each x , where again C_x is the centralizer of x in G . We recall once more that $[G:Z] \geq 4$. As G is the disjoint union of the cosets of Z , it follows that $\mu(Z) \leq 1/4$. (Note that Z is closed and hence measurable.) Now we notice that if $x \in Z$, then $C_x = G$ and so $\mu(C_x) = 1$; on the other hand, if $x \in G - Z$, then C_x has index at least 2 in G , whence $\mu(C_x) \leq 1/2$. Therefore we have

$$\begin{aligned}
\Pr(G) &= \mu \times \mu(C) = \int_G \mu(C_x) d\mu(x) \\
&= \int_Z \mu(C_x) d\mu(x) + \int_{G-Z} \mu(C_x) d\mu(x) \\
&\leq \mu(Z) \cdot 1 + \mu(G-Z) \cdot 1/2 = \mu(Z) + 1/2 - \mu(Z)/2 \leq 5/8.
\end{aligned}$$

3. Further remarks. Let us now return to the case of finite groups. Here the formula $\Pr(G) = k/n$ may be used to good advantage in calculating bounds on $\Pr(G)$ for special classes of groups. For example, the reader may use the class formula to show that for nonabelian p -groups G , $\Pr(G) \leq (p^2 + p - 1)/p^3$. Some information may also be gathered from the theory of group characters [2]. One makes use of the fact that the number of irreducible complex characters of G is just k , together with the fact that $|G| = [G:G'] + n_1^2 + \cdots + n_s^2$, where G' is the commutator subgroup of G , and n_1, \dots, n_s are the degrees of the nonlinear irreducible characters.

Here are some problems for the reader to try:

- (i) $\Pr(G \times H) = \Pr(G) \cdot \Pr(H)$.
- (ii) If $\Pr(G) = 5/8$, then G is nilpotent.
- (iii) If G is finite and $\Pr(G) = 5/8$, then G is the direct product of an abelian group and a 2-group H such that $|H| \geq 8$, H is directly indecomposable and $\Pr(H) = 5/8$.
- (iv) Characterize the groups H having the properties in (iii). (See Miller [6], where the groups with $[G:Z] = 4$ are classified.)
- (v) Derive the bound $\Pr(G) \leq 5/8$ for finite groups by use of the facts from character theory given above.
- (vi) If G is simple and nonabelian, then $\Pr(G) \leq 1/12$, with equality for the alternating group on five letters. (This problem was first posed by J. Dixon.)
- (vii) Study the probabilistic properties of finite groups in general. Some starting points might be Erdős and Turán [4] and Dixon [3]. Of particular interest is a conjecture of Dixon: The probability that two elements chosen at random from a finite simple group G generate G tends uniformly to one as the order of G tends to infinity.

Finally, we would like to encourage the study of a more difficult problem: find lower bounds for $\Pr(G)$. While it is easy to see that no universal lower bound exists, Erdős and Turán [4] have shown that $\Pr(G) \geq (\log_2 \log_2 |G|)/|G|$. C. Ayoub [1] has developed some lower bounds for p -groups of small order.

I am pleased to acknowledge useful conversations with M. Zorn, P. Halmos, and W. Moran. I am also grateful to R. MacKenzie who read the original manuscript.

References

1. C. Ayoub, On the number of conjugate classes in a group, Proc. Internat. Conf. Theory of Groups, Gordon and Breach, New York, 1967, pp. 7-10.
2. C. Curtis and I. Reiner, Representation Theory of Finite Groups and Associative Algebras, Interscience, New York, 1962.

3. J. Dixon, The probability of generating the symmetric group, *Math. Z.*, 110(1969) 199–205.
4. P. Erdős and P. Turán, On some problems of a statistical group-theory, IV, *Acta Math. Acad. Sci. Hung.*, 19 (1968) 413–435.
5. P. Halmos, *Measure Theory*, Van Nostrand, Princeton, N. J. 1950.
6. G.A. Miller, Relative number of non-invariant operators in a group, *Proc. Nat. Acad. Sci. USA*, 30 (1944) 25–28.
7. W. Scott, *Group Theory*, Prentice-Hall, Englewood Cliffs, N. J. 1964.

REMARKS ON THE BESSEL POLYNOMIALS

C. W. BARNES, University of Mississippi

1. Introduction. The Bessel polynomial $y_n(x)$ is defined by Krall and Frink [4] to be the polynomial of degree n , with constant term equal to unity, which satisfies the differential equation

$$(1) \quad x^2 y'' + 2(x+1)y' - n(n+1)y = 0.$$

Krall and Frink discussed the Bessel polynomials from the standpoint of recurrence relations, orthogonality, generating functions, and related matters. Their algebraic properties were considered by Grosswald [2].

In the present note we establish a new result concerning the zeros of the Bessel polynomials. Using a test of Wall [7], which the Bessel polynomials fit in a very natural way, we prove that their zeros have negative real parts. We also give a new proof of a theorem of Dickinson [1], section 6, that the origin is a limit point of zeros of the Bessel polynomials. Our proof of Dickinson's theorem is somewhat simpler than that given in [1] inasmuch as it depends mainly on an application of the maximum modulus principle for analytic functions.

Finally we comment on the history of the Bessel polynomials, and relate them to work of Olds [5] based on Hermite [3].

2. The zeros of the Bessel polynomials. The differential equation (1) is satisfied by

$$(2) \quad y_n(x) = \sum_{k=0}^n \frac{(n+k)!}{(n-k)!k!} \left(\frac{x}{2}\right)^k.$$

These polynomials satisfy the recurrence relations

$$(3) \quad y_{n+1}(x) = (2n+1)xy_n(x) + y_{n-1}(x),$$

where $y_0(x) = 1$, $y_1(x) = 1+x$.

Krall and Frink [4] showed that

$$(4) \quad x^2 y'_n(x) = (nx-1)y_n(x) + y_{n-1}(x)$$

and

$$(5) \quad x^2 y'_{n-1}(x) = y_n(x) - (nx + 1)y_{n-1}(x).$$

Next we require a

LEMMA (Wall [7]). Let $P(x) = x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_n$ be a polynomial with real coefficients, and let $Q(x) = a_1x^{n-1} + a_3x^{n-3} + a_5x^{n-5} + \cdots$ be the polynomial obtained from $P(x)$ by dropping out the first, third, fifth, \cdots terms. Then all zeros of $P(x)$ have negative real parts if and only if the rational function $Q(x)/P(x)$ has a continued fraction expansion of the form

$$\cfrac{1}{c_1x + 1 + \cfrac{1}{c_2x + \cfrac{1}{c_3x + \cfrac{1}{\ddots + \cfrac{1}{c_nx}}}}},$$

where the coefficients c_1, c_2, \cdots, c_n are all positive.

Hence to test a polynomial $P(x)$, we apply the Euclidean algorithm for the greatest common divisor to the polynomials $P(x)$ and $Q(x)$. In the event the sequence of quotients has the form $c_1x + 1, c_2x, \cdots, c_nx$, where each of c_1, c_2, \cdots, c_n is positive, then the zeros of $P(x)$ have negative real parts.

THEOREM 1. The zeros of the Bessel polynomials have negative real parts.

Proof. From (3) we obtain

$$(6) \quad y_{n+1}(-x) = (2n + 1)(-x)y_n(-x) + y_{n-1}(-x).$$

Let $Q_n(x)$ be the polynomials of degree $n-1$ obtained from $y_n(x)$ by dropping out the first, third, fifth, \cdots terms. Then

$$(7) \quad Q_n(x) = \frac{1}{2}\{y_n(x) + (-1)^{n+1}y_n(-x)\}.$$

We can show that

$$(8) \quad Q_{n+1}(x) = (2n + 1)xQ_n(x) + Q_{n-1}(x).$$

For we note that $Q_1(x) = 1$, $Q_2(x) = 3x$, and (8) holds when $n = 2$. We make the inductive hypothesis that (8) holds for all integers m , $2 \leq m \leq n$. Now consider

$$(2n + 1)xQ_n(x) + Q_{n-1}(x)$$

$$\begin{aligned}
&= (2n+1)x\{\tfrac{1}{2}(y_n(x) + (-1)^{n+1}y_n(-x))\} + \tfrac{1}{2}\{y_{n-1}(x) + (-1)^n y_{n-1}(x)\} \\
&= \tfrac{1}{2}\{(2n+1)xy_n(x) + y_{n-1}(x)\} + \tfrac{1}{2}\{(2n+1)x(-1)^{n+1}y_n(-x) + (-1)^n y_{n-1}(-x)\} \\
&= \tfrac{1}{2}y_{n+1}(x) + \tfrac{1}{2}\{(2n+1)x(-1)(-1)^n y_n(-x) + (-1)^n y_{n-1}(-x)\} \\
&= \tfrac{1}{2}y_{n+1}(x) + \tfrac{1}{2}(-1)^n\{(2n+1)(-x)y_n(-x) + y_{n-1}(-x)\} \\
&= \tfrac{1}{2}\{y_{n+1}(x) + (-1)^{n+2}y_{n+1}(-x)\} = Q_{n+1}(x),
\end{aligned}$$

where we used (6) and (7).

Next we show that

$$(9) \quad \frac{y_n(x)}{Q_n(x)} = 1 + x + \cfrac{1 \cdot}{3x + \cfrac{1}{5x + \cfrac{\cdot}{\ddots + \cfrac{1}{(2n-1)x}}}}.$$

We use induction and the recurrence relation (8). Since $y_1(x)/Q_1(x) = 1 + x$, we do indeed have a basis for induction. Thus we suppose that (9) is valid for all integers $2, 3, \dots, n$, and consider the continued fraction

$$1 + x + \cfrac{1}{3x + \cfrac{1}{5x + \cfrac{\cdot}{\ddots + \cfrac{1}{(2n+1)x}}}}.$$

Suppose that $R(x)$ and $S(x)$ denote the numerator and denominator of the continued fraction. Then by the standard recurrence relations for the numerators and denominators of the approximants to a continued fraction we have

$$R(x) = (2n+1)xy_n(x) + y_{n-1}(x),$$

and

$$S(x) = (2n+1)xQ_n(x) + Q_{n-1}(x),$$

since by the inductive hypothesis, $y_n(x)$ is the numerator and $Q_n(x)$ is the denominator of the continued fraction

$$1 + x + \frac{1}{3x + \frac{1}{\ddots + \frac{1}{(2n-1)x}}},$$

and $y_{n-1}(x)$ is the numerator and $Q_{n-1}(x)$ is the denominator of the last approximant to

$$1 + x + \frac{1}{3x + \frac{1}{\ddots + \frac{1}{(2n-3)x}}}.$$

Hence from (3) and (8) it follows that

$$R(x) = y_{n+1}(x) \text{ and } S(x) = Q_{n+1}(x).$$

Thus the expansion (9) is valid for every positive integer n . Theorem 1 now follows directly from the lemma.

Grosswald [2] proved that the zeros of the Bessel polynomials are simple; all zeros are inside the unit circle except for the zero of $y_1(x)$ which is on the unit circle. Theorem 1 improves this last result. It is also established in [2] that for n even, $y_n(x)$ has no real zeros and that for n odd, $y_n(x)$ has a single real, negative zero; finally, if x_n is the real zero of $y_n(x)$ for odd n , we have

$$-1 = x_1 < x_3 < \cdots < x_{n-2} < x_n < \cdots < 0.$$

We can now show that there is an x_n arbitrarily near the origin. Thus we have (see [1], page 954):

THEOREM 2. *The origin is a limit point of the zeros of the Bessel polynomials.*

Proof. Suppose there is a positive number $c < 1$ such that the circle $|z| = c$ contains no zero of the Bessel polynomials. By a classic theorem of Cauchy we have

$$\frac{1}{2\pi i} \int_{|z|=c} \frac{y'_n(z)}{y_n(z)} dz = 0 \quad \text{for } n = 1, 2, 3, \dots.$$

Hence by (4) we have

$$\frac{1}{2\pi i} \int_{|z|=c} \left(\frac{nz-1}{z^2} + \frac{y_{n-1}(z)}{z^2 y_n(z)} \right) dz = 0.$$

Thus

$$\frac{1}{2\pi i} \int_{|z|=c} \frac{y_{n-1}(z)}{z^2 y_n(z)} dz = -\frac{1}{2\pi i} \int_{|z|=c} \frac{nz-1}{z^2} dz = -n.$$

Therefore

$$-n = \frac{1}{2\pi i} \int_{|z|=c} \frac{y_{n-1}(z)}{z^2 y_n(z)} dz,$$

and, by the fundamental estimate on an integral, we have

$$n \leq \frac{1}{c} \max_{|z|=c} \left| \frac{y_{n-1}(z)}{y_n(z)} \right|,$$

or

$$(10) \quad c \leq \frac{1}{n} \max_{|z|=c} \left| \frac{y_{n-1}(z)}{y_n(z)} \right|.$$

Using (5) we have $\{y'_{n-1}(z)/y_{n-1}(z)\} = \{y_n(z)/z^2 y_{n-1}(z)\} - \{(nz+1)/z^2\}$. Therefore, since we shall have

$$\frac{1}{2\pi i} \int_{|z|=c} \frac{y'_{n-1}(z)}{y_{n-1}(z)} dz = 0,$$

it follows that

$$\frac{1}{2\pi i} \int_{|z|=c} \frac{y_n(z)}{z^2 y_{n-1}(z)} dz = \frac{1}{2\pi i} \int_{|z|=c} \frac{nz+1}{z^2} dz = n.$$

As before, we obtain $n \leq 1/c \max_{|z|=c} |y_n(z)/y_{n-1}(z)|$, or

$$(11) \quad c \leq \frac{1}{n} \max_{|z|=c} \left| \frac{y_n(z)}{y_{n-1}(z)} \right|.$$

Considering the estimates (10) and (11), it now follows by the principle of the maximum and minimum modulus that

$$c \leq \frac{1}{n} \left| \frac{y_{n-1}(z)}{y_n(z)} \right|$$

for each point z such that $|z| = c$. In particular, when $z = c$ we have

$$c \leq \frac{1}{n} \left| \frac{y_{n-1}(c)}{y_n(c)} \right| < \frac{1}{n},$$

since for $c > 0$, $y_{n-1}(c) < y_n(c)$. Therefore $c < 1/n$ for every positive integer n . This implies $c = 0$; hence there can be no circle about the origin which is free of zeros of $y_n(x)$, and Theorem 2 is established.

3. Conclusion. Grosswald [2] comments in section 2 on the property of a Bessel polynomial to approximate an exponential and remarks earlier that the polynomials $A_n(x)$ defined by

$$(-1)^{n-1}A_n(x) = (1 - D)^{-n-1}x^n = x^n y_n(2/x),$$

when D is the symbol of derivation, have been used in the proofs of the transcendence of e . These polynomials $A_n(x)$ are discussed in Siegel [6].

In what follows we can be more specific about the connection between an exponential and the Bessel polynomials. This relation also puts into evidence the polynomials $Q_n(x)$ of Wall's test fraction [7].

Olds [5] gave a development of the simple continued fraction for e based on ideas in Hermite's paper [3] in which the transcendence of e was established. Olds obtained two sequences of polynomials $\{T_n(x)\}$ and $\{Z_n(x)\}$ such that

$$T_n(x) = (2n-1)xT_{n-1}(x) + T_{n-2}(x),$$

$$Z_n(x) = (2n-1)xZ_{n-1}(x) + Z_{n-2}(x).$$

In particular,

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_2(x) = 3x^2 + 1, \quad T_3(x) = 15x^3 + 6x,$$

$$Z_0(x) = 1, \quad Z_1(x) = 1, \quad Z_2(x) = 3x, \quad Z_3(x) = 15x^2 + 1.$$

Hence for $n = 1, 2, 3, \dots$ it is immediate that $Z_n(x) = Q_n(x)$ in the notation of section 2, and $T_n(x) + Z_n(x) = y_n(x)$, the Bessel polynomial of degree n .

As a consequence of the fundamental recurrence relations for the numerator and denominators of the approximants of a continued fraction, it follows that the rational functions $T_n(x)/Z_n(x)$, $n = 1, 2, 3, \dots$ are the approximants to the simple continued fraction

$$x + \frac{1}{3x + \frac{1}{5x + \frac{1}{\ddots}}}$$

Olds verified that

$$\lim_{n \rightarrow \infty} \left| \frac{e^{2/x} + 1}{e^{2/x} - 1} - \frac{T_n(x)}{Z_n(x)} \right| = 0,$$

and hence that

$$\frac{e^{2/x} + 1}{e^{2/x} - 1} = x + \frac{1}{3x + \frac{1}{5x + \ddots}}$$

Thus by (9) we have

THEOREM 3. *For $n = 1, 2, 3, \dots$ the sequence of rational functions $\{y_n(x)/Q_n(x)\}$, $x \neq 0$, is the sequence of approximants to the continued fraction*

$$1 + x + \frac{1}{3x + \frac{1}{5x + \ddots + \frac{1}{(2n-1)x + \ddots}}}$$

that is, to the continued fraction for $2e^{2/x}/(e^{2/x} - 1)$.

By the classic theory concerning the approximants to a continued fraction, it is now evident why the Bessel polynomials provide good approximations to an exponential.

References

1. David Dickinson, On Lommel and Bessel Polynomials, *Proc. Amer. Math. Soc.*, 5 (1954) 946-956
2. Emil Grosswald, On Some Algebraic Properties of the Bessel Polynomials, *Trans. Amer. Math. Soc.*, 71 (1951) 197-210.
3. C. Hermite, *Compt. Rend. Acad. Sci. Paris*, 77 (1873) 18-24, 74-79, 285-293.
4. H. L. Krall and Orrin Frink, A New Class of Orthogonal Polynomials: the Bessel Polynomials, *Trans. Amer. Math. Soc.*, 65 (1949) 100-115.
5. C. D. Olds, The simple continued fraction expansion of e , this *MONTHLY*, 77 (1970) 968-974.
6. Carl Ludwig Siegel, *Transcendental Numbers*, *Annals of Mathematics Studies*, Number 16. Princeton University Press, 1949.
7. H. S. Wall, Polynomials whose zeros have negative real parts, this *MONTHLY*, 52 (1945) 308-322.

A MICRONOTE ON A FUNCTIONAL EQUATION

H. N. SHAPIRO, Courant Institute

1. Introduction. It is well known that the only locally integrable (i.e., integrable over every finite interval) solution of the functional equation

$$(1.1) \quad f(x+y) = f(x) + f(y)$$

is of the form $f(x) = cx$, c a constant. In this note we propose to give a very short proof of this.

2. The proof. On the basis of (1.1) and the hypothesis of local integrability one easily verifies the identity

$$(2.1) \quad yf(x) = \int_0^{x+y} f(u) du - \int_0^x f(u) du - \int_0^y f(u) du.$$

Since the right side of (2.1) is invariant under the interchange of x and y , it follows that $xf(y) = yf(x)$. Thus for $x \neq 0$, $f(x)x^{-1} = c$ a constant, or $f(x) = cx$. Since (1.1) implies $f(0) = 0$ this also holds for $x = 0$.

3. Concluding remarks. It is known that the above result remains valid under the weaker hypothesis that $f(x)$ is measurable [1]. Note also that (2.1) asserts that $yf(x)$ is a coboundary.

Reference

1. W. Sierpinski, Sur l'équation fonctionnelle $f(x+y) = f(x) + f(y)$, Fundamenta Math., 1 (1920) 116-122.

AN ADDENDUM TO THE PAPER
"A CHARACTERIZATION OF THE $n \times n$ MATRICES
OVER A FINITE FIELD"

J. V. BRAWLEY, Clemson University and L. CARLITZ, Duke University

In [1], the authors showed that (except for two trivial cases) a ring R has the property that every function f from R to R can be represented by a generalized polynomial [see 1] if and only if for some n and some finite F , R is the ring of $n \times n$ matrices over F . In the present note we obtain an extension of that result. It will be seen that the ideas used here are only slight generalizations of those used in [1].

DEFINITION. Let R be a ring. A *polynomial* in m variables x_1, x_2, \dots, x_m over R is a finite sum of terms of the form

$$a_0^e x_1^{n_1} a_1 y^{n_2} \dots a_{k-1} z^{n_k} a_k^{e'},$$

where x, y, \dots, z can be any of x_1, \dots, x_m , where ε and ε' are in $\{0, 1\}$, and where $k \geq 0$ and the $n_i \geq 1$ are integers.

It is clear that any such polynomial $f(x_1, \dots, x_n)$ defines via substitution a function f from R^m to R in which case the polynomial is said to represent the function f .

THEOREM. *Let $R \neq (0)$ be a ring and let $m \geq 2$ be an integer. Every function from R^m to R , is representable by a polynomial in m variables iff $R = (F)_n$ for some n and some finite field F .*

Proof. Let P denote all functions from R^m to R which are representable by polynomials, and assume $P = R^{R^m}$. R is finite by an argument similar to Lemma 1 of [1]. Also, R has only trivial ideals. To see this, assume $0 \subset I \subset R$ is proper. Select $a \in I - \{0\}$ and $b \in R - I$, let $p: R^m \rightarrow R$ be the function

$$p(r_1, \dots, r_m) = \begin{cases} b; & r_1 = a, r_2 = r_3 = \dots = r_m = 0 \\ 0; & \text{otherwise,} \end{cases}$$

and let $p(x_1, \dots, x_m)$ be the polynomial which represents p . Also, let $\bar{p}(x_1, \dots, x_m)$ be the polynomial over R/I obtained by replacing any coefficient a in $p(x_1, \dots, x_m)$ by $\bar{a} = a + I$. Then \bar{p} defines a map from $(R/I)^m$ to (R/I) and moreover,

$$\bar{p}(\bar{r}_1, \bar{r}_2, \dots, \bar{r}_m) = \overline{p(r_1, r_2, \dots, r_m)}.$$

Thus, $\bar{p}(\bar{a}, \bar{0}, \bar{0}, \dots, \bar{0}) = \bar{p}(\bar{0}, \bar{0}, \dots, \bar{0})$ implies $\overline{p(a, 0, \dots, 0)} = \overline{p(0, 0, \dots, 0)}$ or $\bar{b} = \bar{0}$, a contradiction. Thus, R is a finite simple ring or else $R^2 = (0)$. To show the latter situation is impossible, assume $R^2 = (0)$. Then $(R, +)$ is a finite simple group and thus has prime order p . Since $R^2 = (0)$, only polynomials of the form

$$\varepsilon_0 a_0 + \varepsilon_1 x_1 + \dots + \varepsilon_m x_m,$$

where $\varepsilon_i \in \{0, 1, \dots, p-1\}$ need be considered as representing functions in R^{R^m} . Hence

$$|P| = p^{m+1} = |R^{R^m}| = p^{p^m}$$

which is only valid when $p = 2$ and $m = 1$, a contradiction to the fact that $m \geq 2$. Thus, R is simple so that $R = (F)_n$ for some n and some finite field F .

Conversely assume $R = (F)_n$, and let $f: R^m \rightarrow R$. By Theorem 1 of [1], there exists a polynomial in one variable $p(x)$ which represents the function

$$p(r) = \begin{cases} 1; & r = 0 \\ 0; & r \neq 0. \end{cases}$$

The m -variable polynomial function

$$\sum_{r_m \in R} \dots \sum_{r_1 \in R} f(r_1, \dots, r_m) p(x_1 - r_1) p(x_2 - r_2) \dots p(x_m - r_m)$$

is easily seen to represent f .

Reference

1. J. V. Brawley and L. Carlitz, A characterization of the $n \times n$ matrices over a finite field, this MONTHLY, 80 (1973) 670–672.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

EXPLORING A PLANET

L. FEJES TÓTH, Hungarian Academy of Sciences, Budapest

The surface of a planet can be explored by setting up on the planet a certain number n of bases, serving as starting-points for various operations. The problem of the most economical distribution of the bases may be formulated as follows: How should n bases (points) be distributed on a sphere so as to minimize the greatest distance between a point of the sphere and the base nearest to it? The only values of n for which the solution of this problem is known are $n = 2, 3, 4, 5, 6, 7, 10, 12$ and 14 [1, 2, 3]; for arbitrary values of n there is not even a reasonable conjecture concerning the solution.

Another possibility to explore the planet is to make a set of photographs from a certain number n of satellites guided around the planet. Supposing that the planet hovers in space without rotation and the satellites orbit the planet at the same constant altitude, the problem of the most economical choice of the paths of the satellites is as follows: How should n great circles be distributed on a sphere so as to minimize the greatest distance between a point of the sphere and the great circle nearest to it?

In contrast to the previous problem, there is a chance to solve this problem in its full generality. Indeed, it may be conjectured that the solution is given by n great circles all passing through two antipodal points so as to divide the sphere in $2n$ equal digons. This is trivial for $n = 2$ and it has been proved by Miss V. Rosta [4] for $n = 3$.

The problem is equivalent with its following dual counterpart: How should n points be distributed on a sphere so as to minimize the greatest distance between a great circle and the point nearest to it? The dual counterpart of the above con-

ture says that the points together with their antipodes must be the vertices of a regular $2n$ -gon. This problem may be interpreted as the problem of the most efficient distribution of n observation posts set up by the inhabitants of a planet for the purpose of detecting any satellite guided around the planet by intelligent aliens.

The problem may be rephrased also as follows: Prove or disprove the conjecture that if n equal zones cover the sphere then their width is at least π/n . Here a zone of width w is defined as the parallel domain of a great circle of distance $w/2$.

In this formulation we are at once led to the following generalization: Prove that the total width of any set of zones covering the sphere is at least π .

The problem can be further generalized. Instead of the whole sphere we can consider a convex spherical domain, i.e., a domain any two points of which can be joined by an arc of a great circle lying in the domain. Then we can ask: Is it true that if a convex spherical domain is covered with a set of zones of total width w then it can be covered with one zone of width w ? This problem is a spherical analogue of the well-known "plank problem" of Tarski, first solved by Bang [5, 6] (see also [7, 8, 9, 10, 11, 12, 13]).

Another spherical analogue of the plank problem, unsolved as yet, was proposed by the author ([14], p. 156) more than twenty years ago: Is it true that if a convex spherical domain is covered with a set of digons of total area T then it can be covered with one digon of area T ?

References

1. L. Fejes Tóth, Über die Bedeckung einer Kugelfläche durch kongruente Kugelkalotten, *Mat. Fiz. Lapok*, 50 (1943) 40–46. (Hungarian with German abstract.)
2. K. Schütte, Überdeckung der Kugel mit höchstens acht Kreisen, *Math. Ann.*, 129 (1955) 181–186.
3. G. Fejes Tóth, Kreisüberdeckungen der Sphäre, *Studia Sci. Math. Hung.*, 4 (1969) 225–247.
4. V. Rosta, An extremal arrangement of three great circles on a sphere, *Mat. Lapok*, 24 (1973) (to appear). (Hungarian with English abstract.)
5. Th. Bang, On covering by parallel-strips, *Mat. Tidsskr. B*, (1950) 49–53.
6. ———, A solution of the "plank problem," *Proc. Amer. Math. Soc.*, 2 (1951) 990–993.
7. W. Fenchel, On Th. Bang's solution of the plank problem, *Mat. Tidsskr. B*, (1951), 49–51.
8. H. G. Eggleston, On triangles circumscribing plane convex sets, *J. London Math. Soc.*, 28 (1953) 36–46.
9. D. Ohmann, Eine Abschätzung für die Dicke bei Überdeckung durch konvexe Körper, *J. reine angew. Math.*, 190 (1952) 125–128.
10. ———, Kurzer Beweis einer Abschätzung für die Breite bei Überdeckung durch konvexe Körper, *Arch. Math.*, 8 (1957) 150–152.
11. Th. Bang, Some remarks on the union of convex bodies, *Tolftte Skand. Mat.-Kong. Lund*, 1953 (1954) 5–11.
12. T.-Y. Lee, J.-S. Lin, K.-C. Tong, M.-Y. Zhang, A solution of Bang's "Plank problem." *J. Chinese Math. Soc.*, 2 (1953) 139–143. (Chinese with English abstract.)
13. N. Bognár, On W. Fenchel's solution of the plank problem, *Acta Math. Acad. Sci. Hung.*, 12 (1961) 269–270.
14. L. Fejes Tóth, Lagerungen in der Ebene, auf der Kugel und im Raum, Zweite Auflage, Springer-Verlag, Berlin-Heidelberg-New York, 1972.

WHAT ARE THE LATIN SQUARE GROUPS?

J. J. CARROLL, G. A. FISHER, Bell Laboratories, Indian Hill, Illinois, and

A. M. ODLYZKO, N. J. A. SLOANE, Bell Laboratories, Murray Hill, New Jersey

1. The problem. The following question has arisen in connection with the diagnosis of faults in sequential machines [1]. Let G be a permutation group acting transitively on the symbols $\{1, \dots, n\}$. When is it possible to find n elements $g_1 = e$ (the identity of G), g_2, \dots, g_n such that for each i the symbols $g_1(i), \dots, g_n(i)$ are distinct? Such a sequence, when it exists, is called a **driving sequence**.

Given any sequence g_1, \dots, g_n of elements of G , consider the square array of size n in which the (k, l) -th entry is $g_k(l)$. It is easily seen that g_1, \dots, g_n is a driving sequence if and only if this is a Latin square.

Conversely, given a Latin square of order n in which the first row is normalized to be $1, 2, \dots, n$, we can construct a set of n permutations $g_1 = e, g_2, \dots, g_n$ acting on $\{1, \dots, n\}$ which are defined by $g_k(l) = (k, l)$ -th entry of the Latin square. Let us call a group which can be generated by such a set of permutations a **Latin square group**, or simply **Latin**.

As an example, the Latin square

1	2	3	4
2	1	4	3
3	4	1	2
4	3	2	1

gives the permutations $e, (12)(34), (13)(24)$ and $(14)(23)$, which generate (in this case are actually equal to) the Klein 4-group.

We now see that a permutation group G contains a driving sequence if and only if it contains a Latin square subgroup. The initial problem becomes: which groups contain a Latin square subgroup? We also ask: what are the Latin square groups? Both of these questions are open.

2. Known results. A **regular group** is a transitive group with the property that no element, apart from the identity, fixes any symbol. Such a group is a regular (or Cayley) representation of a group of order n , and conversely every regular representation of a group of order n gives a regular group.

THEOREM. *A regular group is a Latin square group.*

Proof. A regular group G must contain exactly $n - 1$ permutations with no fixed point. Let $g_1 = e, g_2, \dots, g_n$ be a list of the elements of G . Then g_1, \dots, g_n is a driving sequence.

As corollaries, we find that any Frobenius group contains a Latin square subgroup, and any abelian transitive group is Latin.

It can also be shown that for all m , the alternating group on m symbols contains a Latin square subgroup, and for all $m \neq 3, 4$ the symmetric group on m symbols is Latin.

Two permutation groups acting on n symbols are regarded as **different** permutation groups if and only if no relabelling of the symbols transforms one into the other.

By examining all the different permutation groups on ≤ 7 symbols ([2], [3], [5], [6] for $n \leq 6$, [8] for $n = 7$), and by generating all the Latin square groups on ≤ 6 symbols (using the list in [7]) we observed the following. Of the 37 transitive groups on ≤ 7 symbols, all but three contain Latin subgroups. Those 3 all act on 6 symbols and have orders 12, 24 and 60. Of the 30 transitive groups on ≤ 6 symbols, exactly 15 are Latin. In addition (using [8]) all primitive groups on ≤ 9 symbols, with the exception of the group of order 60 on 6 symbols just mentioned, contain Latin subgroups.

There are 9408 reduced Latin squares of order 6. (A reduced Latin square has its first row and column in lexicographic order.) 7776 of these generate the symmetric group. This suggests the conjecture that the probability of the rows of a Latin square of order n generating the symmetric group approaches 1 as $n \rightarrow \infty$. This is supported by Dixon's theorem [4] that two randomly chosen permutations on n symbols generate the symmetric group with probability approaching $3/4$ as $n \rightarrow \infty$.

References

1. J. J. Carroll, Examination of sequential circuits: A model and a method, Ph.D. Thesis, Dept. of Elect. Engin., Illinois Inst. Technology, Chicago, Illinois, May, 1972.
2. A. Cayley, On the substitution groups for two, three, four, five, six, seven, and eight letters, *Quart. J. Math.*, 25 (1890-1891) 71-88, 137-155.
3. F. N. Cole, Note on the substitution groups of six, seven, and eight letters, *Bull. New York Math. Soc.*, 2 (1893) 184-190.
4. J. D. Dixon, The probability of generating the symmetric group, *Math. Z.*, 110 (1969) 199-205.
5. G. A. Miller, Memoir on the substitution-groups whose degree does not exceed eight, *Amer. J. Math.*, 21 (1899) 288-337.
6. ———, Historical note on the determination of all the permutation groups of low degrees, *Collected Works*, Vol. I, Univ. of Illinois Press, Urbana, Illinois, 1935, pp. 1-9.
7. C. R. Rao, S. K. Mitra, and A. Matthai, *Formulas and Tables for Statistical Work*, Statistical Publishing Society, Calcutta, 1966, p. 193.
8. C. C. Sims, Computational methods in the study of permutation groups, pp. 169-183 of J. Leech, editor, *Computational Problems in Abstract Algebra*, Pergamon Press, Oxford, 1969.

CLASSROOM NOTES

EDITED BY ROBERT GILMER

Material for this Department should be sent to David Roselle, Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803.

GEOMETRIC FIT OF A MONOTONIC CUBIC

W. P. COOKE, University of Wyoming

1. Introduction. The problem considered herein arose during a discussion with G. L. Haynes [4] about the use of a cubic to represent the distribution function for a certain random variable. This paper, however, deals only with the computational problem of the least squares fit of a monotonic cubic constrained to pass through the origin and the point (1,1). It should be noted that this problem, as well as the more difficult one where the only constraint is monotonicity, may be solved by the method of convex programming shown in Hartley, Hocking, and Cooke [3]. The paper is, in fact, a computational illustration of a well-known result in constrained optimization. After observing the feasible region in Figure 1, it will be clear that even an ordinary Lagrange-multiplier approach will solve the problem.

The author feels that there are many characteristics of the "geometric fit" exhibited herein that would make this problem a good classroom example for a course in numerical analysis, statistics, or operations research. Particularly it would serve as a graphic introduction to some of the important concepts in nonlinear programming.

2. Formulation of the problem. It is desired to estimate the parameters a_i , $i = 0, 1, 2, 3$, in the function

$$(2.1) \quad y = a_0 + a_1x + a_2x^2 + a_3x^3,$$

subject to the conditions that its graph pass through (0,0) and (1,1) and that $y' \geq 0$ for $0 \leq x \leq 1$. Data-points (x, y) will be observed for $0 \leq x \leq 1$.

It is clear that $a_0 = 0$ and $a_1 + a_2 + a_3 = 1$. Thus the experimental data is required to estimate only two parameters subject to $y' \geq 0$. The monotonicity condition may be written in the form

$$(2.2) \quad (1, x) \begin{bmatrix} a_1 & a_2 \\ a_2 & 3a_3 \end{bmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix} \geq 0 \text{ for } 0 \leq x \leq 1.$$

Using the fact that the matrix of the quadratic form in (2.2) must be positive semi-definite along with $a_1 + a_2 + a_3 = 1$, some simple manipulation allows the following formulation:

Estimate a_1 and a_2 in the expression

$$(2.3) \quad u = a_1z_1 + a_2z_2$$

subject to the restrictions

$$(2.4) \quad \left\{ \begin{array}{l} a_1 \geq 0 \\ a_1 + a_2 \leq 1 \\ 3a_1^2 + a_2^2 + 3a_1a_2 - 3a_1 \leq 0 \end{array} \right\},$$

where $u = y - x^3$, $z_1 = x - x^3$, and $z_2 = x^2 - x^3$. Note that $z_1 \geq 0$ and $z_2 \geq 0$ since $0 \leq x \leq 1$.

Inequalities (2.4) specify the convex feasible region S in which the desired estimate $\alpha^{*'} = (a_1^*, a_2^*)$ must be located. S is shown in Figure 1.

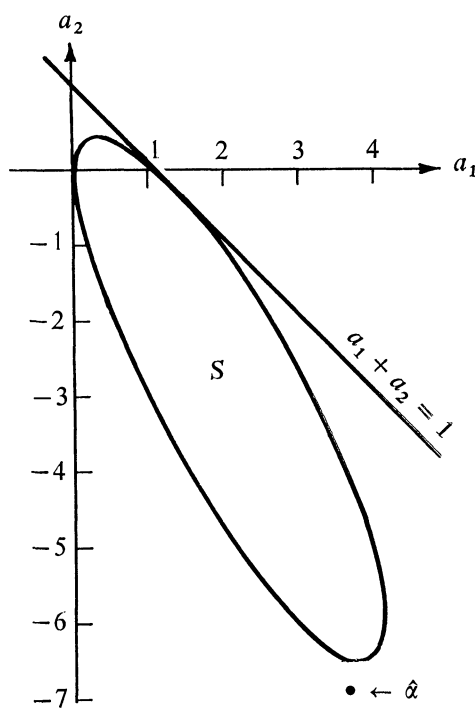


FIG. 1

3. A geometric least squares solution. The criterion for a “best” fit will be that of constrained least squares. Let $u_i^0 = y_i - x_i^3$, $n \geq 2$ be the number of data-points (x_i, y_i) , and let $\alpha' = (a_1, a_2)$. Then, using

$$(3.1) \quad U' = (u_1^0, u_2^0, \dots, u_n^0)$$

and

$$(3.2) \quad Z = \begin{bmatrix} z_{11} & z_{21} \\ z_{12} & z_{22} \\ \vdots & \vdots \\ z_{1n} & z_{2n} \end{bmatrix},$$

the problem is to find α^* so that

$$(3.3) \quad Q(\alpha) = (U - Z\alpha)'(U - Z\alpha)$$

is minimum with $\alpha^* \in S$.

If $\hat{\alpha}$, the estimate of α obtained by the usual least squares method, is in S , then $\hat{\alpha} = \alpha^*$. If not, then using the 2×2 matrix A so that

$$(3.4) \quad A'Z'ZA = I_2$$

and letting $\beta = A^{-1}\alpha = (b_1, b_2)$, it is well known that the point β^* on the boundary of S in β -space (where Q now has circular contours) nearest $\hat{\beta} = A^{-1}\hat{\alpha}$ yields the desired $\alpha^* = A\beta^*$. The argument is found in Graybill [2], in a more general setting in Lewish [5] and Cooke [1], and, more elegantly, in Perlman [6].

It is evident now that if $\hat{\alpha} \notin S$, a geometric solution is obtained by mapping the boundary of S into β -space, finding $\hat{\beta}$, drawing a circle tangent to the boundary of S with $\hat{\beta}$ as center (thus locating β^*), and computing $\alpha^* = A\beta^*$. The reader who is interested in some statistical properties of the estimator α^* is referred to [3].

4. An example. The data in Table 1 will be used to illustrate the necessary computations. This data has been contrived so that $\hat{\alpha} \neq \alpha^*$.

TABLE 1. Data

i	x_i	y_i
0	0	0
1	4/32	1/8
2	5/32	5/8
3	28/32	6/8
4	1	1

TABLE 2. Modified Data

i	u_i^0	z_{1i}	z_{2i}
1	.12305	.12305	.01367
2	.62119	.15244	.02060
3	.08007	.20507	.09570

The data converted to (u, z_1, z_2) -data appears in Table 2. Note that the (x, y) -data when $i = 0$ and $i = 4$ has already been used to yield $a_0 = 0$ and $a_1 + a_2 + a_3 = 1$. Five-decimal accuracy was arbitrarily used.

Using the data from Table 2 the unconstrained least squares estimate of α is $\hat{\alpha}' = (3.68, -6.94)$, with two-decimal accuracy being used for locating $\hat{\alpha}$ in Figure 1.

A quick reference to Figure 1 shows that $\hat{\alpha} \notin S$, so we proceed to the geometric solution for α^* .

The only arithmetic problem is the discovery of the upper-triangular matrix A so that $A'Z'ZA = I_2$, where of course Z is known from the experimental data. Since regardless of the number of data-points $Z'Z$ is a 2×2 matrix, it is not at all difficult to find A .

From the z -data in Table 2,

$$(4.1) \quad A = \begin{pmatrix} 3.52609 & -6.28455 \\ 0 & 20.67397 \end{pmatrix}$$

and $A^{-1} = A'Z'Z$ is

$$(4.2) \quad A^{-1} = \begin{pmatrix} .28360 & .08621 \\ 0 & .04833 \end{pmatrix}.$$

The mapping of the boundary of S to β -space is accomplished by either substituting $A\beta$ for α in (3.3) and sketching the resulting ellipse or mapping point-by-point using $\beta = A^{-1}\alpha$. One should note that while the ellipse in Figure 1 is always the same the mapping will depend on the particular set of experimental data observed. Figure 2 shows the result of the mapping for the data in Table 1 along with the solution for β^* .

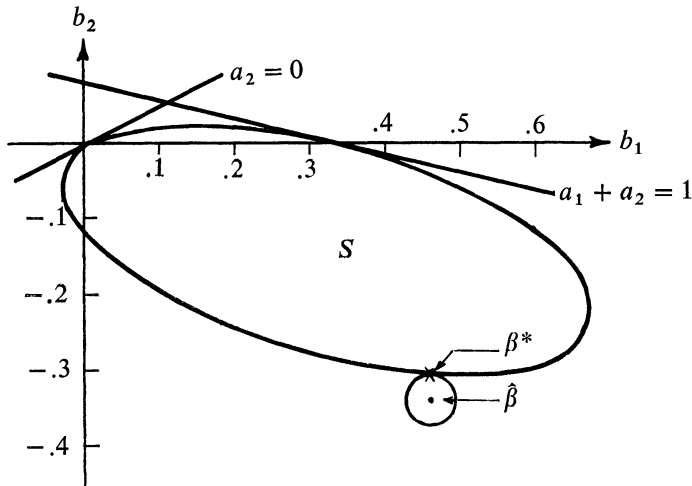


FIG. 2

As well as can be read from Figure 2, $\beta^{*'} = (.45, -.31)$. Then $\alpha^* = A\beta^*$ gives the desired solution

$$(4.3) \quad \alpha^{*'} = (3.54, -6.41).$$

Finally, from $a_1 + a_2 + a_3 = 1$, the estimate $a_3^* = 3.87$ is obtained and the monotonic

cubic is

$$(4.4) \quad y = 3.54x - 6.41x^2 + 3.87x^3.$$

References

1. W. P. Cooke, *Convex Programming Applied to the Estimation of the Parameters of Definite Quadratic Forms and to Related Tests of Hypotheses*, Ph.D. Thesis, Texas A and M University, 1968.
2. F. A. Graybill, *An Introduction to Linear Statistical Models*, McGraw-Hill, New York, 1961 (p. 112).
3. H. O. Hartley, R. R. Hocking, and W. P. Cooke, Least squares fit of definite quadratic forms by convex programming, *Management Science*, 13, No. 11, July (1967) 913-925.
4. G. L. Haynes, Quantification of expertise, The Determination of Empirical and Analytical Spacecraft Parametric Curves, Progress Report II, Part 6, NASA Grant SC-NGR-44-001-027, May, 1966.
5. W. T. Lewish, *Linear Estimation in Convex Parameter Spaces*, Ph.D. Thesis, Iowa State University, 1963.
6. M. D. Perlman, One-Sided Testing Problems in Multivariate Analysis, *The Annals of Mathematical Statistics*, Vol. 40, No. 2, April (1969) 549-567.

A FAMILIAR COMBINATORIAL IDENTITY PROVED BY COMPLEX ANALYSIS

STEVEN MINSKER, Massachusetts Institute of Technology

We shall prove that

$$\binom{n}{0}^2 + \binom{n}{1}^2 + \cdots + \binom{n}{n}^2 = \binom{2n}{n}$$

for all non-negative integers n . Let z be a complex variable, let $\Gamma = \{z : |z| = 1\}$, and let l and j be integers. Note that

$$\frac{1}{2\pi i} \int_{\Gamma} \bar{z} z^l \bar{z}^j dz = \begin{cases} 1 & \text{if } l = j \\ 0 & \text{if } l \neq j \end{cases}.$$

Then

$$\frac{1}{2\pi i} \int_{\Gamma} \bar{z}(z+1)^n (\bar{z}+1)^n dz = \binom{n}{0}^2 + \binom{n}{1}^2 + \cdots + \binom{n}{n}^2,$$

as is seen by replacing $(z+1)^n$ and $(\bar{z}+1)^n$ by their binomial expansions and integrating term-by-term. But

$$\begin{aligned} \frac{1}{2\pi i} \int_{\Gamma} \bar{z}(z+1)^n (\bar{z}+1)^n dz &= \frac{1}{2\pi i} \int_{\Gamma} \bar{z}(2+z+\bar{z})^n dz \\ &= \frac{1}{2\pi i} \int_{\Gamma} \frac{\bar{z}(2z+z^2+1)^n}{z^n} dz = \frac{1}{2\pi i} \int_{\Gamma} \frac{(z+1)^{2n}}{z^{n+1}} dz. \end{aligned}$$

By Cauchy's integral formula, this last expression is just the n th derivative of $(z+1)^{2n}/n!$ evaluated at zero. But this is just $(2n)(2n-1)\cdots(n+1)/n!$, or $\binom{2n}{n}$.

What role should the computer play in mathematics education? Is it an instructional aid? Is programming skill as basic as arithmetic? How much programming and numerical analysis should be a part of the required curriculum?

III. Professional activities. How can two-year college faculties maintain their mathematical interests? How can junior members progress toward higher degrees? How can tyc administrators come to recognize that time spent in professional mathematical activity is as important to the life of the department as time spent in the classroom? Is the manpower now available in the large number of young Ph.D.s a threat or a boon to present tyc departments? How can universities better train their graduate students for tyc positions?

We value your additions, comments, opinions, and recommendations. Please send them to the Committee through Mr. Chinn.

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

ASSOCIATE EDITORS: JOSHUA BARLAZ, ERIC S. LANGFORD. COLLABORATING EDITORS: LEONARD CARLITZ, GULBANK D. CHAKERIAN, HASKELL COHEN, S. ASHBY FOOTE, ISRAEL N. HERSTEIN, MURRAY S. KLAMKIN, DANIEL J. KLEITMAN, ROGER C. LYNDON, MARVIN MARCUS, CHRISTOPH NEUGEBAUER, ALBERT WILANSKY, AND UNIVERSITY OF MAINE PROBLEMS GROUP: EARL M. L. BEARD, GEORGE S. CUNNINGHAM, CLAYTON W. DODGE, OSKAR FEICHTINGER, WILLIAM R. GEIGER, RAMESH GUPTA, GARY HAGGARD, PHILIP M. LOCKE, JOHN C. MAIRHUBER, CURTIS S. MORSE, GRATTAN P. MURPHY, EDWARD S. NORTHAM AND WILLIAM L. SOULE, JR.

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before February 28, 1974.

An asterisk () means neither the proposer nor the editors supplied a solution.*

E 2438*. Proposed by Bernardo Recamán S., Colegio San Carlos, Bogotá, Colombia

For each natural number n , let $f(n)$ denote the least perfect square that begins

with the digits of n ; e.g., $f(2) = 25$, $f(4) = 4$, $f(5) = 529$, $f(12) = 121$. Define $g(n) = \max(f(1), \dots, f(n))$. Do there exist arbitrarily long sequences of consecutive integers on which g is constant?

E 2439. *Proposed by Edmund Umberger, Pennsylvania State University*

Let \cdot and \times denote the usual dot and cross product of vectors in 3-space, and let $*$ denote any of the following operations: the usual scalar multiplication of scalar and vector, multiplication of vector and scalar with the obvious interpretation, or ordinary multiplication of scalar and scalar. How many of the 3^n ways of inserting the symbols \cdot , \times , and $*$ between consecutive vectors of the string $v_1 v_2 \cdots v_{n+1}$ will result in meaningful expressions by suitable insertion of parentheses? For example, $v_1 \times v_2 * v_3 \cdot v_4$ can be made meaningful, whereas $v_1 * v_2 \cdot v_3 \cdot v_4$ cannot.

E 2440. *Proposed by R. C. Entringer, University of New Mexico, and D. E. Jackson, Los Alamos Scientific Laboratories*

Does every permutation of the integers $0, 1, \dots, n$ contain an arithmetic progression of at least three terms?

E 2441. *Proposed by Cornelius Groenewoud, Snyder, N.Y. and F. K. Hwang, Bell Telephone Laboratories*

One has n locations and m teams of n players each. Every week, each team is to send one player to each location where a resolvable round-robin tournament is conducted. Show that it is possible to construct an n -week schedule such that every player goes to every location exactly once, and each player plays against every other player on every other team exactly once, whenever n is a prime power which exceeds m .

E 2442. *Proposed by J. C. Hemperly, University of Maryland*

Let $\omega_1, \omega_2, \dots, \omega_n$ denote the n th roots of unity. Evaluate

$$\sum |\omega_i - \omega_j|^{-2},$$

the sum being taken over all distinct i, j .

E 2443. *Proposed by T. M. Apostol, California Institute of Technology*

Let f_1 and f_2 be two linearly independent functions which are continuous on the bounded interval $[a, b]$. Show that for every pair of constants c_1 and c_2 there exists a continuous function h on $[a, b]$ such that

$$\int_a^b h(x) f_1(x) dx = c_1 \quad \text{and} \quad \int_a^b h(x) f_2(x) dx = c_2.$$

SOLUTIONS OF ELEMENTARY PROBLEMS

Divisors of Powers Plus 1

E 2378 [1972, 906]. *Proposed by D. E. Penney, University of Georgia*

Let a, m and n be natural numbers. Evaluate $(a^m + 1, a^n + 1)$.

Solution by Alan Stein, University of Connecticut at Waterbury. Letting $g = (a^m + 1, a^n + 1)$, the solution is

$$\begin{aligned} g &= a^{(m,n)} + 1 && \text{if } mn/(m,n)^2 \text{ is odd,} \\ g &= 1 && \text{if } mn/(m,n)^2 \text{ is even and } a \text{ is even,} \\ g &= 2 && \text{if } mn/(m,n)^2 \text{ is even and } a \text{ is odd.} \end{aligned}$$

We first establish a lemma: If $(m, n) = 1$, $m > n$, and $e, f = \pm 1$, then

$$\begin{aligned} (a^m + e, a^n + f) &= (a^m + e - ef(a^n + f), a^n + f) = (a^m - ef a^n, a^n + f) \\ &= (a^{m-n} - ef, a^n + f) \text{ since } (a^n, a^n \pm 1) = 1. \end{aligned}$$

Note that both $-ef$ and f are -1 only when both e and f are -1 .

When $(m, n) = 1$, repeated application of the lemma produces g equal to either $(a + 1, a + 1)$ or $(a + 1, a - 1)$, whence $g = a + 1$, or $g = 1$, or $g = 2$. If k is odd, then $a + 1$ divides $a^k + 1$; so, if both m and n are odd, then $g = (a + 1, a + 1) = a + 1$. But if k is even, then $a + 1$ divides $a^k - 1$, so $a + 1$ is not a divisor of $a^k + 1$ unless $a = 1$; hence $g = (a + 1, a - 1)$ if either m or n is even. But $(a + 1, a - 1)$ equals 1 if a is even and 2 if a is odd.

Now let $d = (m, n)$, and note that $(a^m + 1, a^n + 1) = ((a^d)^{m/d} + 1, (a^d)^{n/d} + 1)$ with $(m/d, n/d) = 1$. The solution follows.

Also solved by W. F. Buckler, John Coolidge, R. B. Eggleton, M. G. Greening (Australia), L. Kuipers, J. G. Mauldon, Oto Strauch (Czechoslovakia), Guy Torchinelli, and the proposer.

Editorial Note. O. H. Fraser refers to Exercise 590 in Faddeev & Sominskii, *Problems in Higher Algebra*, which asks for $(x^m + a^m, x^n + a^n)$ and incorrectly omits the answer 2 when x and a are both odd and $mn/(m, n)^2$ is even. Coolidge points out that, when k is odd, $a^k + 1 = a^k - (-1)^k$, so that the first case follows from the solution to E2295 [1972, 398]. Four partial solutions were received.

Matrices with $A \geq 0$ and $A^{-1} \geq 0$

E2379 [1972, 1033]. *Proposed by H. Kestelman, University College, London, England*

Find all matrices A such that both A and A^{-1} have all elements real and non-negative.

Solution by Bennett College Team. A and A^{-1} have all elements real and non-negative if and only if each row and each column of A has exactly one positive element and the rest of the elements are zeros.

Proof. Suppose first that the $n \times n$ matrix $A = (a_{ij})$ has exactly one positive element in each row and each column and that all other elements are zeros. Define $B = (b_{ij})$ in the following way: $b_{ij} = 0$ if $a_{ji} = 0$ and $b_{ij} = 1/a_{ji}$ if $a_{ji} \neq 0$. Clearly $B = A^{-1}$, and B has all elements real and nonnegative.

Conversely, let $A = (a_{ij})$ and $B = (b_{ij})$ be $n \times n$ matrices with real nonnegative elements such that $AB = I$, and assume that one row of A has two or more positive elements, say $a_{ir} > 0$ and $a_{is} > 0$, $r \neq s$. Then if $j \neq i$,

$$0 = \sum_{k=1}^n a_{ik}b_{kj} = \cdots + a_{ir}b_{rj} + a_{is}b_{sj} + \cdots.$$

Therefore $b_{rj} = b_{sj} = 0$ for $j \neq i$ which implies that B is singular, a contradiction. Similarly the assumption that a column of A has two or more positive elements leads to a contradiction.

Also solved by D. T. Adams, K. F. Andersen, Young Archimedes, C. M. Bang, Frederick Carty, John Christopher, R. E. DeMarr, Marjorie Fitting, R. A. Gibbs & L. S. Johnson, J. Z. Hearon, Melvin Henriksen, G. A. Heuer, D. M. Jordon, E. S. Lander, Detlef Laugwitz (Germany), Joel Levy, Milan Lustig (Czechoslovakia), Carolyn MacDonald, Andrzej Makowski (Poland), Jack Neems, William Nusslein, John O'Neill, Robert Patenaude, Kenneth Rosen, D. S. Rubin, W. C. Stone, Phil Tracy, M. R. Vitale, R. M. Warten, G. P. Wene, W. W. Williams, and the proposer.

Editorial Comment. Both Makowski and DeMarr point out that the above result is given in [1] and DeMarr also generalizes the result in [2]. Adams demonstrates in [3] the related result that a compact topological group of matrices with nonnegative elements consists entirely of permutation matrices. He comments that this yields the interesting corollary that such a group must be finite, a result which has implications in algebraic topology (see [4]).

References

1. T. A. Brown, M. Juncosa and V. Klee, *Invertibly positive linear operators on spaces of continuous functions*, Math. Ann., 183 (1969) 105–114. MR 42 # 8314.
2. R. E. DeMarr, *Nonnegative matrices with nonnegative inverses*, Proc. Amer. Math. Soc., 35 (1972) 307–308.
3. D. T. Adams, *On Banach lattices and groups of positive matrices*, to appear.
4. D. R. Brown, *On clans of nonnegative matrices*, Proc. Amer. Math. Soc., 15 (1964) 671–674.

Polynomial Roots Having Rational Imaginary Part

E2380 [1972, 1033]. *Proposed by Erwin Just, Bronx Community College*

Let $f(x)$ be an irreducible polynomial of degree at least three with rational coefficients, and suppose that $f(x)$ has precisely two non-real zeros, $z_1 = p + qi$ and $z_2 = p - qi$, where p and q are real. Could q possibly be rational?

I. *Solution by Allen Charnow and Hwa Tang, California State University at Hayward.* Let K be the splitting field of $f(x)$ over Q , the rationals. Let ϕ be an automorphism in the Galois group $G(K, Q)$ such that $\phi(z_1) = r$, where r is a real zero of f . Then

$$\phi(z_1) - \phi(z_2) = \phi(z_1 - z_2) = \phi(2qi) = \phi(2q) \cdot \phi(i).$$

Assume now $q \in Q$. Then $i \in K$, and $\phi(i)^2 = \phi(i^2) = -1$, so $\phi(i) = \pm i$. Thus $\phi(z_2) = r \pm 2qi$. But no zero of f is of the latter form, a contradiction. Hence q cannot be rational.

II. *Solution by J. Ernest Wilkins, Jr., Howard University.* We shall infer a negative answer to the proposed question from the following more general result.

THEOREM. *Let $Q[x]$ be the set of polynomials in x with coefficients in the field Q of rational numbers. If $f(x)$ is in $Q[x]$, is irreducible over Q , and has two complex zeros of the form $p \pm iq$ with real p and $q^2 \in Q$, $q \neq 0$, then there exists an irreducible $d(x)$ in $Q[x]$ and a constant $a_0 \in Q$ such that $f(x) = a_0 d(x + iq)d(x - iq)$, $d(p) = 0$. In particular the degree of $f(x)$ is even.*

Without loss of generality we may assume that $f(x)$ is a monic polynomial and denote its degree by n . Since $q^2 \in Q$, it is clear that $f(x \pm iq) = F(x) \pm iqG(x)$, in which $F(x)$ and $G(x)$ are in $Q[x]$, $F(x)$ is monic, the degree of $F(x)$ is n , and the degree of $G(x)$ is at most $n - 1$. Since $f(p \pm iq) = 0$ and p is real, $F(p) = G(p) = 0$, and so the (monic) greatest common divisor $d(x)$ of $F(x)$ and $G(x)$ has degree at least 1, and at most $n - 1$. Therefore $f(x \pm iq) = d(x)[\phi(x) \pm iq\psi(x)]$, in which $d(x)$, $\phi(x)$ and $\psi(x)$ are all in $Q[x]$, and so $f^2(x) = H(x)K(x)$, in which $H(x) = d(x + iq)d(x - iq)$ is in $Q[x]$ and has degree not less than 2 and not more than $2(n - 1)$, and $K(x) = \{\phi(x + iq) - iq\psi(x + iq)\}\{\phi(x - iq) + iq\psi(x - iq)\}$ is also in $Q[x]$.

Let $L(x)$ be the (monic) greatest common divisor of $f(x)$ and $H(x)$. Since $f(x)$ is irreducible, either $L(x) = 1$ or $L(x) = f(x)$. In the first case, $f(x)$ and $H(x)$ are relatively prime and so there exist polynomials $\alpha(x)$ and $\beta(x)$ in $Q[x]$ such that $\alpha f + \beta H = 1$. Squaring, and using $f^2 = HK$, we see that $H(\alpha^2 k + 2\alpha\beta f + \beta^2 H) = 1$, and this is impossible since the degree of H is at least two. Therefore $L(x) = f(x)$, $H(x)$ is divisible by $f(x)$, and $H(x) = f(x)A(x)$ for some monic $A(x)$ in $Q[x]$, implying $f^2(x) = f(x)A(x)K(x)$, and so $f(x) = A(x)K(x)$. Since $f(x)$ is irreducible, either $A(x) = f(x)$, in which case $H(x) = f^2(x)$, an impossibility since the degree of H is at most $2(n - 2)$ and the degree of f^2 is $2n$, or $A(x) = 1$, in which case $f = H$. This is the desired conclusion since the reducibility of $d(x)$ obviously implies that of f .

Returning now to the original question, let p_k ($k = 1, 2, \dots, n - 2$) be the zeros of $f(x)$ in addition to z_1 and z_2 . Since $f(x)$ is irreducible these zeros are distinct from each other and from z_1 and z_2 , and for each k , either $p_k - iq$ or $p_k + iq$ is a zero of $d(x)$ since $f(x) = a_0 d(x + iq)d(x - iq)$ and $a_0 \neq 0$. By hypothesis p_k

is real and so both $p_k - iq$ and $p_k + iq$ are zeros of $d(x)$, which therefore has at least $2(n-2) + 1$ zeros, namely p , and $p_k \pm iq$, for $k = 1, 2, \dots, n-2$. Since the degree of $d(x)$ is $n/2$, it follows that $n/2 \geq 2n-3$, implying $n \leq 2$, contrary to the hypothesis that $n \geq 3$. We infer that it is not possible for q^2 to be rational and, *a fortiori*, it is impossible for q to be rational.

Also solved by Robert Gilmer, H. K. Schmidt (Germany), J. H. Smith, and the proposer. Partial solution by Michael Goldberg.

(A, i) Fails in $C[0, 1]$

E 2381 [1972, 1035]. Proposed by E. S. Langford, University of Maine

Suppose that $\{f_n\}$ is a sequence of continuous real-valued functions defined on $[0, 1]$ such that $f_1(x) \geq f_2(x) \geq \dots \geq 0$ for all $x \in [0, 1]$. Suppose further that the only continuous function f such that $f_n(x) \geq f(x) \geq 0$ for all $x \in [0, 1]$ and all $n = 1, 2, \dots$ is the zero function. Is it necessarily true that

$$\int_0^1 f_n(x) dx \rightarrow 0 \text{ as } n \rightarrow \infty?$$

I. Solution by Norman Wilson (graduate student), University of Pittsburgh. In B. R. Gelbaum and J. M. H. Olmstead, *Counterexamples in Analysis* (Holden-Day, San Francisco, 1964, p. 106) there is exhibited a monotonically decreasing sequence $\{f_n\}$ of continuous functions which converges pointwise to the characteristic function χ_C of a Cantor set C of positive Lebesgue measure $m(C)$ (*op. cit.*, pp. 88–89). This provides a counterexample to the assertion, for by the Bounded Convergence Theorem,

$$\int_0^1 f_n(x) dx \rightarrow \int_0^1 \chi_C dm = m(C) > 0.$$

Yet suppose that f is continuous and that $f_n(x) \geq f(x) \geq 0$ for every $x \in [0, 1]$ and every $n = 1, 2, \dots$. If $f(x_0)$ were positive for some $x_0 \in [0, 1]$, then by continuity, f would be strictly positive on some open subinterval I of $[0, 1]$. That is, for every $x \in I$, $0 < f(x) \leq \inf_n f_n(x) = \lim_n f_n(x) = \chi_C(x)$ implying that the Cantor set C contains an interval, a contradiction since C is nowhere dense.

II. Solution by H. Kestelman, University College, London, England. The answer is no. First cover the rational points of $[0, 1]$ by a sequence of open intervals whose lengths sum to $\frac{1}{2}$. Let K denote the subset of $[0, 1]$ not covered by these intervals and for $x \in [0, 1]$ and $n = 1, 2, \dots$ set $f_n(x) = (1 - d(x, K))^n$, where $d(x, K)$ denotes the distance of x to the set K . (Note that K is a closed nowhere-dense set.—Ed.) Then each f_n is continuous, the sequence $\{f_n\}$ is monotonically decreasing, and for every $n = 1, 2, \dots$

$$\int_0^1 f_n(x) dx \geq \frac{1}{2}$$

since $f_n(x) = 1$ for every $x \in K$, and K has Lebesgue measure $\geq \frac{1}{2}$. Suppose that f is continuous and that $0 \leq f(q) \leq f_n(q)$ for every n and every rational $q \in [0, 1]$. Then $0 \leq f(q) \leq \lim_n f_n(q) = 0$ since $1 - d(q, K) < 1$ for all rational q . It follows from the continuity of f that necessarily $f(x) = 0$ for all $x \in [0, 1]$.

III. *Comment by the proposer.* The space $C[0, 1]$ of continuous real-valued functions on $[0, 1]$ is a Riesz space under the pointwise operations and under the integral (L^1) norm this space enjoys all of the properties of an abstract L -space except for completeness. It is known that in an abstract L -space, whenever $\{f_n\}$ is a decreasing sequence of positive elements, then $f_n \rightarrow 0$ in the order sense if and only if $f_n \rightarrow 0$ in norm. The problem asks if this is still true in $C[0, 1]$, and the answer is no as shown above.

Professors W. A. J. Luxemburg and A. C. Zaanen have kindly pointed out that this problem appears as Exercise 18.14 (p. 104) of their book, *Riesz Spaces* (Vol. I), North Holland, Amsterdam, 1971. The result was also noted in the same authors' earlier paper, *Notes on Banach function spaces* (X), Nederl. Akad. Wetensch. Proc. Ser. A67 = Indag. Math. 26(1964), 493–506. In general, a normed Riesz space is said to have Property (A, i) by these authors if $\|f_n\| \rightarrow 0$ whenever $f_n \downarrow 0$.

Also solved by John Annulis, Sheldon Axler, Anders Bager (Denmark), Fred Barber, Gerald Beer, T. A. Bick, Robert Breusch, Benjamin Burrell, Frederick Carty, R. C. Detmer, Harley Flanders, Leon Gerber, Gary Gundersen, M. L. J. Hautus (Netherlands), Ellen Hertz, Lee Hill, Lee Keener, J. A. Kelingos, P. G. Kirmser, E. M. Klein, Detlef Laugwitz (Germany), Joel Levy, John MacBain, Glenn Meyers, L. F. Meyers, Paul Milnes, S. E. Minear, William Nuesslein, John Oman & A. L. Perrie, M. Bhaskara Rao (England), Jürg Rätz (Switzerland), The Ronald Reagan Memorial Problems Group of CSU (Hayward), L. A. Ringenberg, Kenneth Rosen, Sullivan Ross, T. Šalát (Czechoslovakia), Nan-Shan Shou, Michael Steiner (Sweden), Manfred Stoll, John Swetits, U. R. U., R. M. Warten, M. Weiss & C. Wexler & M. Goldstein, J. K. Yates, and the proposer. Eight (incorrect) affirmative answers to the question were received. Most attempted to use Lebesgue's Dominated Convergence Theorem, sometimes in conjunction with Egoroff's Theorem. Two incorrect counterexamples were received.

Editorial Comment. All of the solutions submitted depended on the existence of a closed nowhere-dense subset C of $[0, 1]$ of positive Lebesgue measure. In all circumstances, the subset presented was either the classical Cantor set as used by Wilson or the subset presented by Kestelman constructed by deleting open intervals about the rationals (or any countable dense subset.) Once the set C was found, the construction of the monotonically decreasing sequence $\{f_n\}$ of continuous functions converging pointwise to χ_C proceeded in basically one of four directions: (1) Citing the result from Gelbaum and Olmstead; (2) Constructing f_n in a piecewise linear fashion; (3) Using a $d(x, C)$ construction as in Kestelman's solution; (4) Noting that C is closed so that it is a G_δ set and hence there exists a sequence $G_1 \supseteq G_2 \supseteq \dots$ of open sets such that $\bigcap G_n = C$. Urysohn's Lemma was then used to show the existence of continuous functions g_n such that $g_n(x) = 1$ if $x \in C$ and $g_n(x) = 0$ if $x \in [0, 1] \setminus G_n$, and f_n was then taken to be either $g_1 g_2 \dots g_n$ or $g_1 \wedge g_2 \wedge \dots \wedge g_n$.

Swetits comments that if $\{f_n\}$ is a monotonically decreasing sequence of continuous functions satisfying the conditions of the problem and if f is the pointwise limit of the f_n , then necessarily $f(x_0) = 0$ at every point of continuity x_0 of f . It follows that if f is Riemann-integrable, then necessarily $f = 0$ almost everywhere so that $\int f(x) dx = 0$. This means that there can be no counterexample with a Riemann-integrable limit function. (Note that if C is a closed nowhere-dense set of positive

measure, then the set of discontinuities of the characteristic function χ_C is precisely C — a set of positive measure.)

On a lighter note, Kelings remarks that “Operating on the premise that any pathological situation in real variables that mortal man can conjecture is possible, the answer is no.”

Shades of E 712

E 2382 [1972, 1034]. *Proposed by Thomas Hughes, Fort Worth, Texas*

One has a number of balls, identical in appearance; one of the balls is known to be slightly heavy, another slightly light by the same amount, and the rest have a standard weight. It is desired to isolate both the light and heavy balls, using only three weighings on a “triple platform balance.” (A triple platform balance consists of three arms forming a Y , equally spaced at intervals of 120° ; these are supported at the center, and at the end of each arm is a pan. If n balls are placed in each of the three pans, then one can tell whether each of the three sets of balls is heavier, lighter, or the same weight as n standard balls; note however that the heavy ball and the light ball weigh as much as two standard balls.)

What is the largest number of balls from which one can identify both the heavy ball and the light ball in only three weighings?

Composite solution edited from those given by Guy Torchinelli, SUNY at Buffalo, and O. P. Lossers, Technological University, Eindhoven, The Netherlands. As a result of the first weighing, the set of balls is divided into four subsets, namely the three subsets formed by balls weighed in the three pans and the subset of unweighed balls. As a result of the second weighing, each of these four sets is divided into four subsets. If the first two weighings balance, then each of these sixteen sets has standard weight. There are 13 possible outcomes for the last weighing: equilibrium $(0, 0, 0, 0)$, and 12 permutations of one pan heavy and one pan light $(+, -, 0, 0)$. The third weighing, then, cannot locate the bad balls among the 16 subsets, when the first two weighings both yield equilibrium, if:

- (1) One of the sets contains at least 5 balls (since there are $5 \cdot 4 = 20$ possible choices for the bad balls).
- (2) One set contains 4 balls and another set at least 2 balls ($4 \cdot 3 + 2 = 14$ choices).
- (3) Each of two sets contains 3 balls (even though there are only $3 \cdot 2 + 3 \cdot 2 = 12$ choices).
- (4) One set contains 3 balls and each of four sets contains 2 balls ($3 \cdot 2 + 4 \cdot 2 = 14$ choices).
- (5) Each of seven sets contains 2 balls (14 choices).

When 23 or more balls are separated into sixteen subsets, there are 7 extra balls over and above one ball per subset. Thus one of the five indeterminate situations occurs, so three weighings will not suffice for 23 balls.

We now show it possible to make the decision in 3 weighings if there are 22 balls. Consider the array below where the set in row i and column j is weighed in pan i in the first weighing and in pan j in the second weighing ("pan 4" contains those balls set aside).

<i>Weighing 2</i>				
<i>Weighing 1</i>	Pan 1	Pan 2	Pan 3	Pan 4
Pan 1	1, 2	3, 4	5	6
Pan 2	7	8, 9	10, 11	12
Pan 3	13, 14	15	16, 17	18
Pan 4	19	20	21	22

If each of the first two weighings results in a permutation of $(+, -, 0, 0)$, then we are left with at most 4 balls from which we are to find the lighter and the heavier in the final weighing, and this is trivial. If the first weighing results in $(0, 0, 0, 0)$ and the second in $(+, -, 0, 0)$, for example, then in the same row of the above array the heavy ball is in the first column and the light ball is in the second column. It is a simple matter to check that a final weighing of $\{1, 7, 8\}$, $\{2, 9, 15\}$, $\{3, 13, 19\}$, $\{4, 14, 20\}$ in the four pans locates the bad balls. Note that the last column, though not of the same type as the others, may be changed, when necessary, by adding to it balls known to be of the correct weight. Finally, if the first two weighings both balance, then the bad balls are one of the six pairs found in an intersection of a row and a column. They are easily located in the third weighing of $\{1, 10, 16\}$, $\{3, 11, 13\}$, $\{8, 14, 17\}$, $\{2, 4, 9\}$. It follows that the required decision can be made for any number of balls not exceeding 22.

Also solved by Jordi Dou (Spain), H. S. Hahn, Peter Klein (Sweden), H. L. Nelson, Kenneth Rosen, and the proposer.

Editorial Note. Ten balls was the most common of the 7 incorrect answers received, and 64 the largest. Two solvers report that two weighings suffice for 9 balls, but, curiously, not for 8. For 9 balls, weigh $\{1, 2\}$, $\{3, 4\}$, $\{5, 6\}$, $\{7, 8, 9\}$ in the first weighing. If $(0, 0, 0, 0)$, then weigh $\{1, 7\}$, $\{3, 8\}$, $\{5, 9\}$, $\{2, 4, 6\}$ the second time; if $(+, 0, 0, -)$ the first weighing, then weigh $\{1, G\}$, $\{2, 7\}$, $\{8, G\}$, $\{9\}$ the second time, where G denotes any known good ball. The other cases are easy. For 8 balls, which *can* be done in two weighings, eliminate ball 9 on the first weighing above. If $(0, 0, 0, 0)$, then use $\{1, 5\}$, $\{2, 3\}$, $\{4, 7\}$, $\{6, 8\}$ in the second weighing. Other cases are trivial.

With one weighing, four balls is the maximum.

Walter Bluger notes the similarity of this problem to problem E 712 [1947, 46] for a two-pan balance.

A general solution for four or more weighings is solicited (see E 712, for example). By the subset argument above, three weighings of equilibrium produce 64 subsets, so $64 + 6 = 72$ balls is an upper bound for four weighings. Other considerations seem to indicate a smaller least upper bound. The editors have a solution for 60 balls in four weighings, copies of which can be obtained by writing to the Problems Group.

Catalan Numbers

E2383 [1972, 1034]. *Proposed by E. T. Ordman, University of Kentucky*

Let n be a nonnegative integer. For $p = 1, 2, \dots$ define

$$S_p(n) = \sum_{k=0}^{\lfloor n/2 \rfloor} \left[\binom{n}{k} - \binom{n}{k-1} \right]^p,$$

where we make the usual conventions regarding binomial coefficients. It is easy to evaluate $S_1(n)$. Evaluate $S_2(n)$.

Solution by Richard Gibbs and Harold Stocker, Fort Lewis College, Durango, Colorado. Let

$$T(n) = \sum_{k=0}^{n+1} \left[\binom{n}{k} - \binom{n}{k-1} \right]^2.$$

By the symmetry of the binomial coefficients, $T(n) = 2S_2(n)$. Now

$$T(n) = \sum_{k=0}^{n+1} \left[\binom{n}{k}^2 - 2\binom{n}{k} \binom{n}{k-1} + \binom{n}{k-1}^2 \right].$$

From the identity $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k}$ we obtain

$$2\binom{n}{k} \binom{n}{k-1} = \binom{n+1}{k}^2 - \binom{n}{k}^2 - \binom{n}{k-1}^2.$$

Therefore,

$$T(n) = 2 \sum_{k=0}^{n+1} \binom{n}{k}^2 + 2 \sum_{k=0}^{n+1} \binom{n}{k-1}^2 - \sum_{k=0}^{n+1} \binom{n+1}{k}^2.$$

Using the identity $\sum_{j=0}^m \binom{m}{j}^2 = \binom{2m}{m}$ and the fact that $\binom{n}{-1} = \binom{n}{n+1} = 0$ we obtain

$$\begin{aligned} T(n) &= 4 \binom{2n}{n} - \binom{2n+2}{n+1} = 4 \binom{2n}{n} - 2 \left[\binom{2n}{n-1} + \binom{2n}{n} \right] \\ &= 2 \binom{2n}{n} - 2 \binom{2n}{n-1}. \end{aligned}$$

Therefore

$$S_2(n) = \binom{2n}{n} - \binom{2n}{n-1} = \frac{1}{n+1} \binom{2n}{n}.$$

Thus $S_2(n)$ is just the n th Catalan number.

Also solved by Günter Bach (Germany), Anders Bager (Denmark), M. T. Bird, D. M. Bloom, Dieter Bode (Germany), Frederick Carty, H. W. Gould, M. G. Greening (Australia), Robert Heller, J. D. Hiscocks, O. P. Lossers (Netherlands), Alexandru Lupas (Romania), Milan Lustig (Czechoslovakia), Kumer Murty & Ram Murty, M. R. Railkar (India), Jürg Rätz (Switzerland), Kenneth Rosen, F. C. Smith, Phil Tracy, E. Trost (Switzerland), David Zeitlin, and the proposer.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers — The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before February 28, 1974.

An asterisk () means neither the proposer nor the editors supplied a solution.*

5934. *Proposed by R. C. Buck, University of Wisconsin*

What entire functions f obey the inequality

$$|f(z+a)| \leq |f(z)| |f(a)|$$

for all z and for two non-zero choices of a whose ratio $\beta = a_2/a_1$ is irrational?

5935*. *Proposed by K. Selucký, Brno, Czechoslovakia*

Suppose

$$\frac{1}{x_1} + \frac{1}{s-x_2} + \frac{1}{s-x_3} + \frac{1}{x_4} = \frac{1}{s-x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \frac{1}{s-x_4},$$

where $s > x_1 \geq x_2 \geq x_3 \geq x_4 \geq \frac{1}{2}s$. Prove $x_1 = x_2$ and $x_3 = x_4$.

5936. *Proposed by David Styer, University of Cincinnati*

Is there a function f , bounded and holomorphic in $|z| < 1$, and a polynomial p such that, in $|z| < 1$, f/p assumes every complex value infinitely often with at most one exception?

5937. *Proposed by Albert Wilansky, University of Reading, England*

Give an example of a vector space with two comparable (unequal) norms such that it is barreled with the larger norm and a Banach space with the smaller. [A barreled (locally convex topological vector) space is one such that every closed absolutely convex absorbing set is a neighborhood of 0.]

5938. *Proposed by Detlef Laugwitz, Nieder-Ramstadt, Germany*

Let G be an archimedean subfield of the ordered field F , and let A be an order preserving automorphism of F . Is it true that $AG = G$?

5939. *Proposed by J. W. Andrushkiw, Seton Hall University*

Let $f(z) = a_0 + a_1z + \cdots + a_nz^n$ be a polynomial with real positive coefficients. Show that if for some k , $0 \leq k \leq n-3$, the inequality $a_k a_{k+3} \geq 3a_{k+1}a_{k+2}$ holds, then $f(z)$ cannot have all of its zeros located in the complex left halfplane.

SOLUTIONS OF ADVANCED PROBLEMS

Metrizability from a Neighborhood Base

5859 [1972, 524]. *Proposed by L. A. Feldman, Stanislaus State College, California*

Prove that a T_0 topological space (X, T) is a metric space if and only if each $x \in X$ has a neighborhood base of open sets

$$\{B_r(x) \mid r \in (0, 1]\}$$

such that (1) if $r, s \in [0, 1]$ with $r \leq s$, and $B_0(x) = \{x\}$ then $B_r(x) \subset B_s(x)$; (2) if $B_r(x) \cap B_s(y) \neq \emptyset$ for $r, s \in [0, 1]$, where $0 < r + s \leq 1$, then for some t where $0 < t < r + s$, we have $x \in B_t(y)$.

Solution by D. G. Belanger, University of South Alabama. We first show that X is T_2 . Let $\{x_i\}$ be a net approaching distinct points x and y . From property (2), $x \in B_t(y)$ and $y \in B_t(x)$ for every $t \in (0, 1]$, hence, because X is T_0 , $x = y$. Let

$$d(x, y) = \min(1, \inf \{r \mid y \in B_r(x)\}).$$

Since X is T_2 , $d(x, y) = 0$ if and only if $x = y$. To prove symmetry assume that $d(x, y) = r$, $d(y, x) = s$ and $r + \varepsilon < s$ for some $x, y \in X$, $\varepsilon > 0$. Since $B_{r+\varepsilon}(x) \cap B_0(y) \supset \{y\}$, $x \in B_{r+\varepsilon}(y)$ contradicting the assumption. Thus $d(x, y) = d(y, x)$. If $y \in B_r(x)$ and $y \in B_s(z)$ then $z \in B_t(x)$ where $0 < t < r + s$ by property (2). This proves the triangle inequality $d(x, z) \leq d(x, y) + d(z, y)$. Since $B_r(x) = \{y \mid d(x, y) < r\}$, the metric d induces the topology on X .

Also solved by R. A. Christiansen, David Singmaster (England), P. van der Steen (Netherlands), R. K. Tomaki, and the proposer.

Note. Tomaki and Christiansen point out that this result is a consequence of the metrization theorem of A. H. Stone to be found on p. 196 of Dugundji, *Topology*.

Submodules as Direct Summands

5862 [1972, 667]. *Proposed by R. C. Wagner, Fairleigh Dickinson University*

A submodule N of the R -module M is said to be *pure* if for every $r \in R$, $rN = N \cap rM$. Prove that if R is a commutative Noetherian ring with unit and M is a finitely generated R -module for which every submodule is pure, then every submodule is a direct summand of M .

Solution by James R. Smith, Appalachian State University. Suppose some submodule of M is not a direct summand. M is Noetherian since it is a finitely generated module over a Noetherian ring. So let N be a maximal element in the set of all submodules of M which are not direct summands. Then, if $x \in M - N$, we must have $Rx \cap N \neq 0$, for, otherwise, N would be a summand since $N + Rx$ is a

summand. Let $x \in M - N$. Then $N + Rx = M$, for $N + Rx$ is a summand of M , so if P is a complementary summand for $N + Rx$ and $y \in P$, by the above if $y \neq 0$ we have $Ry \cap N \neq 0$. But $P \cap N = 0$, so $y = 0$. Now let $z \in M - N$ be such that $N : z$ ($\{a \in R \mid az \in N\}$) is maximal in the set of all $N : y$ for $y \in M - N$. Let $N : z = A$ and suppose that A is generated by a_1, a_2, \dots, a_m . We prove by induction that there is a $z' \in M - N$ such that $a_i z' = 0$ for $i = 1, 2, \dots, n$ and $N : z' = A$. For $n = 1$ we have $a_1 z \in N$, so $a_1 z = a_1 w$ for some $w \in N$ since N is pure. Thus $a_1(z - w) = 0$ and $z - w \in M - N$, and $N : (z - w) \supset N : z$, so $N : (z - w) = N : z = A$. Also $a_1(z - w) = 0$. Suppose true for $n < m$. Let w be such that $N : w = A$ and $a_i w = 0$ for $1 \leq i \leq n$. Then $Ra_{n+1}w$ is a pure submodule of M and $a_{n+1}w \in Ra_{n+1}w$, so $a_{n+1}w = a_{n+1}(ra_{n+1}w)$ for some $r \in R$. Let $z' = w - ra_{n+1}w$. Then $a_1 z' = 0$ for $i = 1, 2, \dots, n + 1$, $z' \in M - N$ and $N : z' = A$. So we have that this is true for m and there is some $w \in M - N$ such that $N : w = A$ and $A \cdot w = 0$. But then $R \cdot w \cap N = 0$, a contradiction to an earlier observation.

Also solved by Jim Brewer & Phil Montgomery & Ed Rutter, S. H. Cox, D. Z. Djoković, Bruce Ferrero, E. R. Gentile (Argentina), Robert Gilmer, A. A. Jagers (Netherlands), Daniel Opitz, T. G. Parker, and the proposer.

Irreducibles in Integral Domains

5863 [1972, 667]. *Proposed by P. R. Chernoff, University of California, Berkeley*

Let D be an integral domain with infinitely many elements. Assume that every non-unit in D has an irreducible factor.

Prove that D has infinitely many irreducibles or infinitely many units.

Solution by Bruce Ferrero, Princeton University. If D has no irreducibles, then every non-zero element must be a unit, and D is an infinite field. If D has an irreducible, but only finitely many, let P be their product. For each $n \geq 1$, let $A_n = P^n + 1$. Then A_n must be a unit, since any irreducible divides P . If D also has only finitely many units, then $A_n = A_m$ for some $n > m$, and hence $P^{n-m} = 1$, which is absurd.

Also solved by James Alonso, D. D. Anderson, J. T. Arnold, Jim Brewer & Phil Montgomery & Ed Rutter, P. K. Garlick, S. N. Gersten, Robert Gilmer, G. S. Glazer, A. A. Jagers (Netherlands), Wells Johnson, L. E. Mattics, Barbara Osofsky, J. R. Smith, Hwa Tang, W. C. Waterhouse, and the proposer.

On the Parts in a Partition

5864 [1972, 668]. *Proposed by G. E. Andrews, Pennsylvania State University*

Let P_n denote the set of partitions of n into positive integers. For each $\pi \in P_n$, let $d(\pi)$ denote the number of different parts of π , and let $\#(\pi)$ denote the total number of parts of π . Prove that for $n \geq 1$,

$$\sum_{\pi \in P_n} (-1)^{\#(\pi)} 2^{d(\pi)} = \begin{cases} 2(-1)^n & \text{if } n \text{ is a square,} \\ 0 & \text{otherwise.} \end{cases}$$

Solution by Harry Lass, Jet Propulsion Laboratory, California Institute of Technology. Let $n = n_\pi = \sum_{k=1}^n k\alpha_k$ be a representation of n into positive integers, and define $U(0) = 0$, $U(\alpha_i) = 1$ for α_i a positive integer. The number of distinct parts of n_π is $d(\pi) = \sum_{i=1}^n U(\alpha_i)$, and the total number of parts of n_π is $\#(\pi) = \sum_{i=1}^n \alpha_i$.

We note that

$$\sum_{\alpha=0}^{\infty} (-1)^{\alpha} 2^{U(\alpha)} s^{k\alpha} = \frac{1-s^k}{1+s^k}, \quad |s| < 1.$$

If we let

$$f(n) = \sum_{\pi} (-1)^{\#(\pi)} 2^{d(\pi)}, \quad f(0) = 0,$$

with π ranging over all partitions of n into positive integers, then

$$F(s) = \sum_{n=0}^{\infty} f(n)s^n$$

$$\prod_{r=1}^{\infty} \frac{1-s^r}{1+s^r} = \prod_{r=1}^{\infty} (1-s^{2r})(1-s^{2r-1})^2.$$

A theorem of Jacobi states that

$$\prod_{r=1}^{\infty} (1-s^{2r})(1+s^{2r-1}z)(1+s^{2r-1}z^{-1}) = \sum_{n=-\infty}^{\infty} s^{n^2} z^n$$

for $|s| < 1$, $z \neq 0$. Setting $z = -1$ yields

$$F(s) = \sum_{n=-\infty}^{\infty} (-1)^n s^{n^2} = 1 + 2 \sum_{n=1}^{\infty} (-1)^n s^{n^2},$$

so that $f(n) = 2(-1)^n$ if n is a square, $f(n) = 0$ otherwise.

Also solved by G. E. Andrews, L. Carlitz, N. J. Fine, M. G. Greening (Australia), L. E. Mattics, M. R. Modak (India), David Newman, Allen Stenger, Phil Tracy, and by the proposer.

Editorial Note. G. E. Andrews points out that the formula may be found in a paper by J.W.L. Glaisher, Proc. London Royal Soc., 24(1875-6) Formula III, p. 252.

A Non-Archimedean Vector Lattice

5868 [1972, 780]. Proposed by B. C. Anderson, Henry Ford Community College

Show that the following theorem becomes false if "Archimedean" is omitted. R^n is an Archimedean vector lattice with respect to the order generated by a cone K if and only if there are n linearly independent vectors $\mathbf{v}^{(k)}$ such that $K = \{x = (x_j) \in R^n : \sum_{j=1}^n x_j \mathbf{v}_j^{(k)} \geq 0; k = 1, 2, \dots, n\}$. (Note: the word "Archimedean" is inadvertently omitted on p. 9 of A.L. Peressini, *Ordered Topological Vector Spaces*.)

Solution by J. T. Annulis, University of Arkansas, Monticello. Consider R^2 with the lexicographic order, i.e., with positive cone $K = \{(x, y): x > 0 \text{ or } x = 0 \text{ and } y \geq 0\}$. It is easily verified that (R^2, K) is a vector lattice which is not Archimedean. Suppose there exists $v^{(1)} = (v_1^{(1)}, v_2^{(1)})$ and $v^{(2)} = (v_1^{(2)}, v_2^{(2)})$ such that

$$K = \{x = (x_j) \in R^2: x_1 v_1^{(i)} + x_2 v_2^{(i)} \geq 0 \text{ for } i = 1, 2\}.$$

Then, since $(1, 0)$ and $(0, 1)$ are elements of K , we have $v_j^{(i)} > 0$ for $j = 1, 2$ and $i = 1, 2$. By linear independence both $v_2^{(1)}$ and $v_2^{(2)}$ cannot be zero. We may then assume that $v_2^{(1)} > 0$. Let $x_2 = -2/v_2^{(1)}$. Then if $x = (x_1, x_2)$ where $x_1 = 1$ if $v_1^{(1)} = 0$ and $x_1 = 1/v_1^{(1)}$ if $v_1^{(1)} > 0$, we have (x_1, x_2) is an element of K . But $x_1 v_1^{(1)} + x_2 v_2^{(1)} < 0$, a contradiction to

$$K = \{(x_1, x_2): x_1 v_1^{(i)} + x_2 v_2^{(i)} \geq 0 \text{ for } i = 1, 2\}.$$

Also solved by A. L. Peressini, G. C. Schmidt, and by the proposer.

A Random Inscribed n -gon

5869 [1972, 780]. *Proposed by Anatol Rapoport, University of Toronto*

Let n points be chosen at random on the circumference of a unit circle. Show that the expected area of the inscribed n -gon is given by

$$A(n) = \pi \left[1 - n! \sum_{j=1}^{\infty} \frac{(2\pi)^{2j} (-1)^{j-1}}{(n+2j)!} \right].$$

Solution by J.G. Little and O. G. Ruehr, Michigan Technological University. After choosing the n points, we break the circle at one of them, calling it $\phi_0 = 0$. The remaining $n-1$ points, in order, are distributed as the order statistics $\phi_1, \dots, \phi_{n-1}$ of the uniform distribution on the interval $(0, 2\pi)$. Letting $\phi_n = 2\pi$, we find that the differences $\theta_i = \phi_i - \phi_{i-1}$, $i = 1, \dots, n$ have the common probability density function

$$f(\theta) = \left(\frac{n-1}{2\pi} \right) \left(1 - \frac{\theta}{2\pi} \right)^{n-2}.$$

[Feller, William, *An Introduction to Probability Theory and its Applications*, New York (1966), pp 21-22.]

We triangulate the n -gon by drawing radii through the n points and chords through adjacent points. Since the area of the i th triangle is $\frac{1}{2} \sin \theta_i$, we have

$$\begin{aligned} A(n) &= E \left(\frac{1}{2} \sum_{i=1}^n \sin \theta_i \right) = \frac{1}{2} \sum_{i=1}^n E(\sin \theta_i) \\ &= \frac{n(n-1)}{4\pi} \int_0^{2\pi} \sin \theta \left(1 - \frac{\theta}{2\pi} \right)^{n-2} d\theta. \end{aligned}$$

Two successive partial integrations yield

$$A(n) = \pi \left[1 - \int_0^{2\pi} \sin \theta \left(1 - \frac{\theta}{2\pi} \right)^n d\theta \right].$$

Replace $\sin \theta$ by its MacLaurin series and integrate term by term, employing the familiar beta function integral to obtain

$$A(n) = \pi \left[1 - n! \sum_{j=1}^{\infty} \frac{(-1)^{j-1} (2\pi)^{2j}}{(2j+n)!} \right].$$

Note that the last expression is a (convergent) asymptotic expansion for n large.

Also solved by Günter Bach (Germany), A. A. Jagers (Netherlands), A. F. W. Jägst (Netherlands), Harry Lass, J. G. Mauldon, Thomas Spencer, Luis Verde-Star, and the proposer.

Notes. Lass points out that $\lim_{n \rightarrow \infty} A(n) = \pi$.

The proposer compares $A(n)$ with $A^*(n)$, the designated area of the regular inscribed n -gon, and observes the following:

$$A(n) \sim \pi(1 - (2\pi)^2 / (n+1)(n+2)), \quad A^*(n) \sim \pi(1 - (2\pi)^2 / 6n^2),$$

$$A^*(n) - A(n) \sim 10\pi^3 / 3n^2,$$

$$A(n+1) - A(n) \sim 6\pi^3 / n^3 \sim 9(A^*(n+1) - A^*(n)).$$

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN

with the assistance of the mathematics departments of St. Olaf and Carleton Colleges

COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER, CARLETON COLLEGE

We invite readers to submit reviews of significant recent college-level mathematics books. We especially encourage reviews based on classroom use, or comparative reviews of several related books. Reviews should ordinarily not exceed two pages (per book) typed double spaced. Manuscripts of reviews as well as books submitted for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057.

Introduction to Projective Geometry. By C. R. Wylie, Jr. McGraw-Hill, New York, 1970. vii + 556 pp. \$12.40.

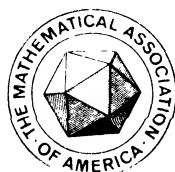
Projective Geometry and Algebraic Structures. By R. J. Mihalek. Academic Press, New York, 1972. xi + 220 pp. \$9.75. (Telegraphic Review, August–September, 1972.)

The appearance of two well-written texts is welcome in any subject, but is doubly so in projective geometry, a field in which there are relatively few books in print. Different parts of the subject are emphasized in these two books. The emphasis in Mihalek is on the axiomatic foundations of the subject and on the relation between the geometrical properties of the plane and the algebraic structure of the system which coordinatizes the plane.

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

THE OFFICIAL JOURNAL OF
THE MATHEMATICAL ASSOCIATION OF AMERICA



VOLUME 80

NUMBER 10

CODEN: AMMYAE

CONTENTS

Number Fifty-two	HARLEY FLANDERS	1099
Prime Numbers and Brownian Motion	PATRICK BILLINGSLEY	1099
Correction to "Unique Factorization Domains"	P. M. COHN	1115
Correction to "A History of the Prime Number Theorem"	L. J. GOLDSTEIN	1115

MATHEMATICAL NOTES

Complements and Comments	ROBERT GILMER AND DAVID ROSELLE	1116
A Problem on Series	G. J. O. JAMESON	1119

RESEARCH PROBLEMS

Monthly Research Problems, 1969-73	R. K. GUY	1120
----------------------------------------------	-----------	------

CLASSROOM NOTES

The Minimal Polynomial of a Linear Transformation	M. D. BURROW	1129
Another Proof of the Rational Decomposition Theorem	H. G. JACOB	1131
A Proof of the Chain Rule for Derivatives in n -space	A. G. FADELL	1134

MATHEMATICAL EDUCATION

Survival of the Two-year College Mathematics Teacher	P. A. LINDSTROM	1135
----------------------------------------------------------------	-----------------	------

ELEMENTARY PROBLEMS AND SOLUTIONS	1138
---------------------------------------------	------

ADVANCED PROBLEMS AND SOLUTIONS	1146
------------------------------------------	------

REVIEWS	1152
-------------------	------

(Continued on inside cover)

DECEMBER

1973

NEWS AND NOTICES	1163
MATHEMATICAL ASSOCIATION OF AMERICA	1164
The Fifty-fourth Summer Meeting of the Association	1164
Report of the Treasurer for the Year 1972	1180
Honorary Life Membership for Professor Emory P. Starke	1181
April Meeting of the Maryland-District of Columbia-Virginia Section	1182
June Meeting of the Northeastern Section	1183
Mathematical Sciences Employment Register — Open Register	1183
Acknowledgement	1184
Calendars of Future Meetings	1186
INDEX	1187

NOTICE TO AUTHORS

Specialized research is usually unsuitable; see *Statement of Policy* (vol. 76, p. 2). Manuscript preparation: Please use the *Manual for Monthly Authors* (vol. 78, p. 1) and follow the format in current issues of the MONTHLY. Manuscripts should be typewritten, triple-spaced with wide margins; submit two copies and keep one for protection against loss.

Backlog: Main Articles 12 months, Math. Notes 13 months, Research Problems 7 months, Classroom Notes 11 months, Math. Education 10 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to ALEX ROSENBERG, Department of Mathematics, Cornell University, Ithaca, N.Y. 14850; NOTES, etc.: to the corresponding Associate Editor; ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N.W., Washington, D.C. 20036.

HARLEY FLANDERS, *Editor*

ALEX ROSENBERG, *Editor-Elect*

ASSOCIATE EDITORS

JOSHUA BARLAZ	J. G. HARVEY	SEYMOUR SCHUSTER
E. R. BERLEKAMP	ERIC S. LANGFORD	J. ARTHUR SEEBACH, JR.
JANE W. DI PAOLA	P. D. LAX	E. P. STARKE
ROBERT GILMER	ARTHUR MATTUCK	LYNN A. STEEN
RICHARD GUY	M. W. POWNALL	JAMES WENDEL
RAOUL HAILPERN	GIAN-CARLO ROTA	

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are \$12.50. For nonmembers the subscription price is \$18.00.

PUBLISHED BY THE ASSOCIATION at Washington, D. C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Washington, D. C., and additional mailing offices.

Copyright © The Mathematical Association of America (Incorporated), 1973

PRINTED IN THE UNITED STATES OF AMERICA

NUMBER FIFTY-TWO

The retiring editorial staff nursed from cradle through publication ten MONTHLYS per year for the last five years, also two Slaughter papers. We hope most have been satisfactory, some excellent. Our debt to our authors and referees is great, also to our readers for their steady support and encouragement.

I personally am grateful to all of my associate and collaborating editors. They worked hard and competently to a high professional standard. They join me in wishing our successors well.

Harley Flanders

PRIME NUMBERS AND BROWNIAN MOTION

PATRICK BILLINGSLEY, The University of Chicago

Because it factors into a product of prime numbers, each integer contains within it a kind of Brownian motion path, and the mathematics of Brownian motion can be used to derive theorems about the factorization. Despite the persistent notion that a result stated in probability language is rather less true than it might otherwise be, I shall state these theorems in probability language and even give them probabilistic proofs. As a matter of fact, there will be little in the way of real proofs, since for the most part I shall only illustrate general results by examples and special cases. For this there is the authority of William Feller, who used to tell us, his students, that the best in mathematics, as in art, letters, and all else — that the best consists of the general embodied in the concrete. Although at first I thought that was simply an antimilitary sentiment, I did eventually understand it as the intellectual-esthetic principle he intended and have tried ever since to keep it at the front of my mind.

The paper has three sections. In Section 1, the mathematical model for a particle in Brownian motion is defined and some of its properties described. Section 2, which provides the link between Brownian motion and primes, concerns random walk: one successively tosses a coin and successively moves along a scale, one unit in the positive

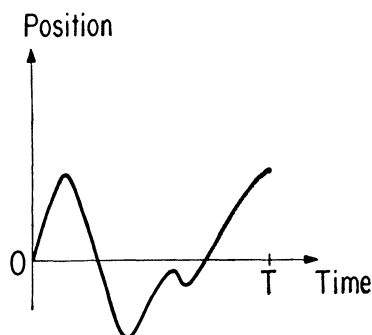
Patrick Billingsley received his Princeton Ph.D under William Feller. Except for a period of Navy service, he has been at the University of Chicago, where he is presently Professor of Mathematics and Statistics. He was a Fulbright lecturer for a year at the University of Copenhagen, and held a Guggenheim fellowship at Cambridge University. Professor Billingsley's main research interest is probability theory, and he is a fellow of the Institute of Mathematical Statistics. His publications include the books: *Statistical Inference for Markov Processes*, 1961, *Ergodic Theory and Information*, 1965, *Convergence of Probability Measures*, 1968, *Elements of Statistical Inference* (with D. L. Huntsberger), 1973. *Editor*.

or negative direction, according as the coin falls heads or tails. Here it is shown how a distant random walk looks approximately like a Brownian motion and how the Brownian motion model therefore leads to limit theorems associated with random walk. Section 3 discusses the random walk which a randomly chosen integer generates through its prime factorization: one successively examines the primes, 2, 3, 5, ..., and successively moves along a scale, one unit in the positive or negative direction, according as the prime appears in the factorization or not. It turns out that, because of the arithmetic fact that distinct primes individually divide an integer if and only if their product does, this factorization random walk has many of the properties of the ordinary coin-tossing random walk; in particular, it too can be approximated by Brownian motion, and it is shown how this leads to limit theorems associated with factorization into primes.

In addition to the elements of real analysis, the paper makes use of statistical concepts such as mean, variance, independence, and Gaussian distribution.

1. Brownian motion. Imagine suspended in a fluid a particle bombarded by molecules in thermal motion. The particle will perform that irregular and seemingly random movement first described by the biologist Robert Brown in 1828. Since we shall be concerned with just one component of this motion, imagine it projected on a vertical axis: At each instant t of time we note the height $x(t)$ of the particle above a fixed horizontal plane. Over T units of time, the motion of the particle, which we take to start at 0, is described by the positions $x(t)$ for $0 \leq t \leq T$ —that is, by a continuous real function x on $[0, T]$ with $x(0) = 0$. This leads us to consider the collection $C_0[0, T]$ of such functions x .

FIGURE 1

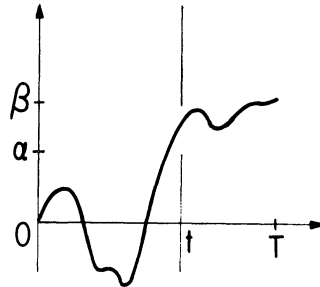


For technical reasons, we make $C_0[0, T]$ into a metric space by taking the distance between two of its elements to be the maximum vertical distance between their graphs. This topology, the uniform topology, is of little direct concern here; it is brought in mostly as evidence that the discussion to follow does have a rigorous basis.

The random motion of the particle is described by an assignment of probabilities

$P_T(A)$ to subsets A of $C_0[0, T]$; $P_T(A)$ represents the chance that the path traced out by the particle lies in A , or is described by a function x that lies in A . Probabilities represent long-run relative frequencies. If the total on a pair of dice is observed, the possible outcomes are 2, 3, ..., 12. If many pairs of balanced dice are rolled independently, the proportion among them producing the outcome 7 will be about $1/6$. If a particle in Brownian motion is observed for T units of time, the possible outcomes are the various elements of $C_0[0, T]$. If many independently moving particles are observed, the proportion among them producing paths that lie in A will be about $P_T(A)$. Although the interpretation of probability involves such multiple observations, in the mathematical theory we speak of a single roll of the dice, the probability the roll produces a 7 being $1/6$; in the same way, we speak of a single particle, the probability it traces out a path that lies in A being $P_T(A)$.

FIGURE 2



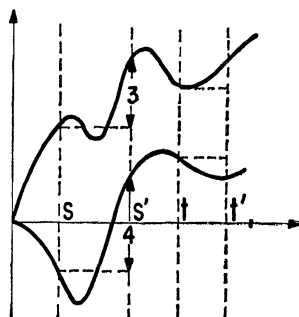
The set $[x: \alpha \leq x(t) \leq \beta]$, consisting of the paths that go through the gate in Figure 2, represents the event that at time t the particle will lie between α and β ; it is assigned probability

$$(1) \quad P_T[x: \alpha \leq x(t) \leq \beta] = \frac{1}{\sqrt{2\pi t}} \int_{\alpha}^{\beta} e^{-u^2/2t} dt.$$

Thus the distribution of the position at time t follows the Gaussian curve with mean 0 and variance t . That the mean is 0 reflects the fact that the particle is as likely to go up as to go down; there is no drift. The variance t grows linearly; this indicates that the particle tends to wander away from its starting point and, having done so, suffers no force tending to restore it to that starting point. The equation (1) can be extended: the increment over $[s, t]$ has a Gaussian distribution with mean 0 and variance $t - s$.

The other important property of Brownian motion is this: Suppose $s < s' < t < t'$, and consider for example the event $A = [x: x(s') - x(s) \geq 3]$ that the particle undergoes an upward displacement of at least 3 units during the time interval $[s, s']$, together with the event $B = [x: x(t') - x(t) < 0]$ that the particle undergoes a

FIGURE 3



downward displacement during the time interval $[t, t']$. The top path in Figure 3 lies in A but not in B , and the bottom path lies both in A and in B . The probabilities of A and B and of their intersection $A \cap B$ are related by

$$(2) \quad P_T(A \cap B) = P_T(A)P_T(B).$$

Thus A and B satisfy the definition of independence; that is, that the displacement the particle undergoes during $[s, s']$ in no way influences the displacement it undergoes during $[t, t']$. This implies a kind of lack of memory. Although the future behavior of the particle depends on its present position, it does not depend on how the particle got there. Equation (2) has a more general form showing that the increments over any number of disjoint intervals are statistically independent of one another.

The equations (1) and (2), together with generalized versions of them, determine all the probabilities $P_T(A)$. (This ignores a technical point: $P_T(A)$ cannot be defined for every subset A of $C_0[0, T]$, but it can for every Borel set A —that is, for every A in the σ -field generated by the sets open in the uniform topology.) It was one of Norbert Wiener's achievements to prove in 1923 that there does exist an assignment of probabilities satisfying these rules, and P_T (the corresponding measure on the Borel sets) is accordingly called Wiener measure. Here we shall take its existence for granted.

Brownian motion, as described by Wiener measure, obeys a transformation law having consequences strange and deep. Suppose that a particle performs a Brownian motion for T units of time, and suppose that, in the function representing its path, we contract the time scale by the factor T and the position scale by the factor \sqrt{T} . According to the law in question, the new path will be exactly like that of a particle that has been in Brownian motion for 1 unit of time.

To understand why, let x and y be the old and new paths, so that x lies in $C_0[0, T]$, y lies in $C_0[0, 1]$, and

$$(3) \quad y(t) = \frac{1}{\sqrt{T}} x(tT), \quad 0 \leq t \leq 1.$$

Of course (3) defines a mapping

$$(4) \quad C_0[0, T] \rightarrow C_0[0, 1].$$

The transformation law says that, if x is a random path in $C_0[0, T]$ distributed according to P_T , then y is a random path in $C_0[0, 1]$ distributed according to P_1 . (Technically, if ϕ_T is the mapping (4), then $P_1 = P_T \phi_T^{-1}$.) Now, according to (1), the quantity $x(tT)$ is a Gaussian random variable with mean 0 and variance tT . Multiplying a Gaussian random variable by a constant a multiplies its mean by a and its variance by a^2 , and the new variable is also Gaussian. Therefore the distribution of $y(t)$ as defined by (3) follows the Gaussian curve with mean 0 and variance t (since $T^{-\frac{1}{2}} \cdot 0 = 0$, $(T^{-\frac{1}{2}})^2 \cdot tT = t$), the first requirement for Brownian motion. Contracting time by the factor T leads from a path over $[0, T]$ to a path over $[0, 1]$, and the vertical rescaling by $1/\sqrt{T}$ makes the variances work out right. Moreover, x has (over disjoint intervals) independent increments, and it is intuitively clear that monotone changes of the time and position scales cannot convert independent increments into dependent ones. So the transformation (3) must preserve the other property of Brownian motion, that of independent increments. This argument, which makes the transformation law plausible, can be converted into a complete proof.

By means of the transformation defined by (3), it is possible to see that, whatever positive values ε and K may have, a Brownian path over $[0, 1]$ will with probability exceeding $1 - \varepsilon$ have somewhere a chord with slope exceeding K . The trick is this: We want, over $[0, 1]$, a Brownian path y with a steep chord. We obtain it not directly, but by applying the transformation (3) to a Brownian path x over $[0, T]$ with T suitably chosen. Choose T so large that x will, with probability exceeding $1 - \varepsilon$, have somewhere a chord with slope exceeding, say, 1. Such a T exists because even the most miraculous event will happen in the long run (the monkeys at the typewriters), and the occurrence of a chord with slope exceeding 1 is a modest miracle indeed. At the same time, choose T to exceed K^2 . If x has a chord with slope exceeding 1, and if x and y are related by (3), then y has a chord with slope exceeding \sqrt{T} , which in turn exceeds K .

Since ε may be taken arbitrarily small and K arbitrarily large, a Brownian path over $[0, 1]$ must with probability 1 have chords with arbitrarily great slope. There must also be chords with arbitrarily large negative slope, and in fact, chords (very short ones) with extreme slopes are dense along the path. In rigorous and more elaborate form, these arguments show that, if A is the set of paths in $C_0[0, 1]$ of unbounded variation, then $P_1(A) = 1$. A path of unbounded variation represents the motion of a particle that in its wanderings back and forth travels an infinite distance, and at this point physicists lose interest because of their obsession with reality. The fact is mathematically interesting, however, and so is the fact that $P_1(A) = 1$ if A is the set of functions in $C_0[0, 1]$ that are nowhere differentiable. Constructing a

continuous, nowhere differentiable function is difficult, but drawing an element from $C_0[0,1]$ randomly according to P_1 produces such a function with probability 1.

In what follows we shall be mainly concerned with sets that correspond more closely with reality. Although Sections 2 and 3 will involve the transformation (3) and T 's that exceed 1, for the rest of this section we shall take $T = 1$. We shall need (1) for the case $T = t = 1$:

$$(5) \quad P_1[x: \alpha \leq x(1) \leq \beta] = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-u^2/2} du.$$

Suppose $\alpha \geq 0$ and consider the event $[x: \max x(t) \geq \alpha]$ that the particle achieves the height α at some time t with $0 \leq t \leq 1$. First,

$$\begin{aligned} P_1[x: \max x(t) \geq \alpha] &= P_1[x: \max x(t) \geq \alpha \text{ and } x(1) \geq \alpha] \\ &\quad + P_1[x: \max x(t) \geq \alpha \text{ and } x(1) < \alpha]. \end{aligned}$$

The two probabilities on the right here can be proved equal, roughly because once the particle achieves the height α it is as likely, in the absence of drift, to wander upward and finish above α at time 1 as to wander downward and finish below α . Thus

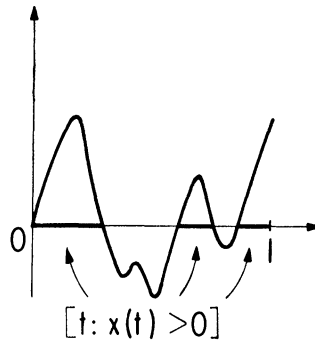
$$P_1[x: \max x(t) \geq \alpha] = 2P_1[x: \max x(t) \geq \alpha \text{ and } x(1) \geq \alpha].$$

Since the condition $\max x(t) \geq \alpha$ is superfluous in the presence of the condition $x(1) \geq \alpha$, the right side here is $2P_1[x: x(1) \geq \alpha]$, and (5) with $\alpha \geq 0$ and $\beta = \infty$ now implies

$$(6) \quad P_1[x: \max x(t) \geq \alpha] = \frac{2}{\sqrt{2\pi}} \int_{\alpha}^{\infty} e^{-u^2/2} du.$$

Thus we have the distribution of the greatest positive excursion.

FIGURE 4



Although to make it rigorous requires some effort, this derivation of (6) has an intuitive appeal. The next result will be stated without any proof, and like many

ex cathedra assertions, it runs counter to intuition. Consider the set $[t: x(t) > 0]$ of time points t , $0 \leq t \leq 1$, for which the particle is above 0. This set is a union of intervals (infinitely many, contrary to Figure 4). Denote by bars the Lebesgue measure of this set, the sum of the lengths of the constituent intervals: $|\bar{[t: x(t) > 0]}|$. The distribution of this quantity, the total time spent above 0, is given by

$$(7) \quad P_1[x: \alpha \leq |\bar{[t: x(t) > 0]}| \leq \beta] = \frac{1}{\pi} \int_{\alpha}^{\beta} \frac{du}{\sqrt{u(1-u)}}$$

for $0 \leq \alpha \leq \beta \leq 1$. This is Paul Lévy's arc sine law, so called because carrying out the integration leads to the arc sine function.

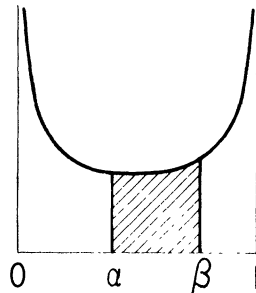


FIGURE 5

Figure 5 shows the shape of the density, the area of the shaded region representing the right side of (7). The curve is U-shaped, so that if the length $\beta - \alpha$ of the interval is fixed, the probability of $\alpha \leq |\bar{[t: x(t) > 0]}| \leq \beta$ grows as the interval nears 0 or 1, being smallest when the interval is centered on $\frac{1}{2}$. This is odd because the time spent above 0 has mean $\frac{1}{2}$ by symmetry, and ordinarily values near the mean of a random quantity are more likely to occur than are values far removed from the mean, whereas here the situation is just the opposite.

For general accounts of Brownian motion, see [4] and [7].

2. Random walk. Imagine a particle moving about at random on the nodes of a cubic lattice. The particle can move in any of six directions (north, south, east, west, up, down) to an adjacent node. The direction is determined by the roll of a balanced die, the particle moves to the next node, and the die is rolled once more to determine the direction of the next move, and so on. Figure 6 shows five steps of

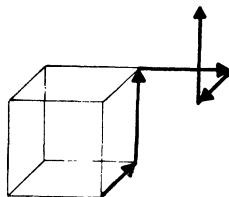
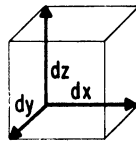


FIGURE 6

such a random walk, together with one of the cells of the cubic lattice. The figure is in

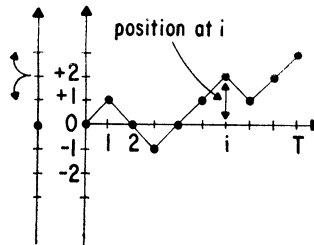
the spirit of a venerable vector analysis book which began a proof of Gauss's theorem by enjoining the reader to consider "an infinitesimal element of volume of dimensions dx , dy , and dz ." This injunction was accompanied by a nicely labelled diagram like Figure 7, which was said to show such an infinitesimal element of volume "much enlarged." Well, Figure 6 is much enlarged too, and if the cubes of the lattice are really very small and the particle moves very rapidly from node to node it is natural to expect the motion to approximate Brownian motion.

FIGURE 7



We shall explore a one-dimensional version of this idea. Consider a vertical axis with the integer points $0, \pm 1, \pm 2, \dots$ marked off on it. We start at 0, toss a coin, and move upward one unit if the coin falls heads and downward one unit if the coin falls tails. In the new position ($+1$ or -1), we toss the coin again and move up or down one unit according as it falls heads or tails, and we continue this way for T steps, T being here an integer. If we take one unit of time to execute each step of this random walk and proceed at a uniform rate from one node to the next, our progress is described by a function like that in Figure 8, a polygonal path whose height over i is the position at i —that is, the position after the i th step. Of the 2^T such paths, each has probability 2^{-T} . (Various aspects of random walk are discussed in [3].)

FIGURE 8



The path can also be viewed as describing the fluctuations in a gambler's fortune. The position on the vertical axis represents the gambler's fortune (relative to his initial capital, so that he starts conventionally at 0), and it moves up or down one unit—say one pound—according as he wins or loses the next play.

The random walk path has some of the properties of a Brownian motion path over $[0, T]$. In the first place, for integers with $i < i' < j < j'$, the displacements undergone over the time intervals $[i, i']$ and $[j, j']$ are independent because they depend on disjoint sets of tosses and the tosses are assumed independent (the coin has no memory). Thus the path has essentially independent increments (for intervals

with nonintegral endpoints the increments can be slightly dependent). The distance moved in one step has mean

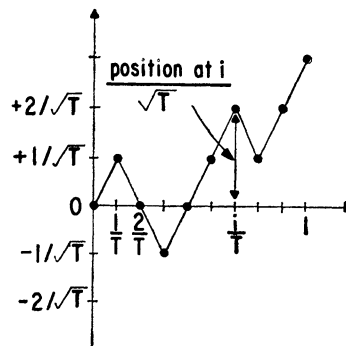
$$(8) \quad (+1)\frac{1}{2} + (-1)\frac{1}{2} = 0$$

and variance

$$(9) \quad (+1)^2\frac{1}{2} + (-1)^2\frac{1}{2} = 1,$$

and so the position at i has mean 0 and by independence has variance i , another property of Brownian motion (see equation (1)). (For nonintegral t , the position at t has mean 0, but the variance is only approximately t .) Although the polygonal

FIGURE 9



character of the path is not shared by Brownian motion, contraction of the two scales will make the straight-line segments in Figure 8 disappear in the limit as $T \rightarrow \infty$.

Suppose we contract the time scale by a factor T and the vertical scale by a factor \sqrt{T} , applying the transformation (3) to pass from Figure 8 to Figure 9. In Figure 8 the segments have length $\sqrt{2}$, whereas in Figure 9 they are very short for large T , having length of the order $1/\sqrt{T}$. If Figure 8 represented a Brownian motion path over $[0, T]$, then, as explained in Section 1, Figure 9 would represent a Brownian motion path over $[0, 1]$. The transformation (3) leaves invariant those characteristics (means, variances, independence of increments) the original path shares with Brownian motion and tends to mask those characteristics (piecewise linearity) it does not share. Thus we can hope that the curve in Figure 9 will be very like a Brownian motion path for large T . And indeed, it is true that

$$(10) \quad \text{Prob}[\text{path} \in A] \rightarrow P_1(A) \quad (T \rightarrow \infty)$$

for subsets A of the space $C_0[0, 1]$, where $P_1(A)$ is Wiener measure. There are 2^T paths like the one in Figure 9, and $\text{Prob}[\text{path} \in A]$ is 2^{-T} times the number of them that lie in A .

For an illustration of this theorem, suppose the A in (10) is the set $[x: \alpha \leq x(1) \leq \beta]$ of paths in $C_0[0, 1]$ that over the point $t = 1$ have a height between α and β . Since the height over $t = 1$ in Figure 9 is $1/\sqrt{T}$ times the position at T in the original

random walk, (10) and (5) together imply

$$(11) \quad \text{Prob} \left[\alpha \leq \frac{\text{position at } T}{\sqrt{T}} \leq \beta \right] \rightarrow \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-u^2/2} du.$$

This is the classical DeMoivre-Laplace central limit theorem for Bernoulli trials. It describes the position after a large number of steps in a random walk, or the gambler's fortune at the end of an evening's play of T ventures. If $-\alpha = \beta = .9$, the limit in (11) is about .6. If $T = 100$, the gambler thus has probability approximately .6 of ending the evening within $.9 \times \sqrt{100} = 9$ pounds of his initial capital.

Suppose now that A is the set in (6), the set of paths in $C_0[0, 1]$ having somewhere a height at least α (here $\alpha \geq 0$). The path in Figure 9 lies in A if at some time during the evening's play the gambler's fortune is at least $\alpha\sqrt{T}$ pounds above his initial capital, and by (10), the probability of this converges to the right side of (6). For $\alpha = 1.7$, the value of this limit is about .1. With $T = 100$, this gives an approximate probability of .1 that the gambler will have been at least $1.7 \times \sqrt{100} = 17$ pounds ahead at the time he should have quit.

Finally, suppose A is the set $[x: \alpha \leq |t: x(t) > 0| \leq \beta]$ in (7). During the evening the gambler is ahead a certain fraction of the time; if the curve in Figure 9 represents the history of his fortunes, it belongs to the set A if and only if this fraction lies between α and β . The chance of this event is by (10) and (7) about equal to the area of the shaded region in Figure 5. If we compute the areas, the chance the gambler is ahead between 45% and 55% of the time turns out to be only about .06, whereas the chance he is ahead more than 90% of the time is about .2. In one evening in five the gambler will thus be ahead more than 90% of that evening's play. By symmetry, in one evening in five the gambler will be ahead less than 10% of that evening's play. To convince him in the first [second] case that his experience is due merely to chance and not to his being Fortune's favorite [Fortune's fool] will be difficult [impossible].

We have applied (10) to three interesting sets A . If A is the set of functions in $C_0[0, 1]$ of unbounded variation, then $P_1(A) = 1$, as explained in Section 1, while $\text{Prob}[\text{path} \in A] = 0$ because the curve in Figure 9 is visibly of bounded variation. Thus (10) fails for certain subsets A of $C_0[0, 1]$. The mathematical fact is that (10) holds for every set (Borel set) A whose boundary ∂A (boundary in the sense of the uniform topology) satisfies $P_1(\partial A) = 0$ —a condition which holds in our three applications but not if A is the set of functions of unbounded variation. A complete proof of this theorem uses a combination of probability theory and functional analysis; the details can be found in [1].

3. Prime divisors. According to the fundamental theorem of arithmetic, each integer has a factorization into primes, a factorization unique except for order (see [5], for example). Let $f(n)$ be the number of distinct primes in the factorization of n ; we do not count multiplicity: $f(3^4 \cdot 5^2)$ is 2, not 6. The table shows some values of the

n	2	3	4	5	6	7	...	29	30	31	...	209	210	211	...
$f(n)$	1	1	1	1	2	1	...	1	3	1	...	2	4	1	...

function f . It rises slowly. The smallest n 's with respective f -values 2, 3, and 4 are $2 \cdot 3 = 6$, $2 \cdot 3 \cdot 5 = 30$, and $2 \cdot 3 \cdot 5 \cdot 7 = 210$. The fact that there are infinitely many primes implies, however, that f assumes arbitrarily large values; since $f(p) = 1$ for prime p , the same fact implies that f infinitely often drops back to 1.

Since f varies in this irregular fashion, it is natural to ask after its average behavior. For example, it can be shown that

$$(12) \quad \frac{1}{N} \sum_{n=1}^N f(n) \approx \log \log N$$

(see the remarks following (17) below). Since $\log \log 10^{70} \approx 5$, the typical integer under 10^{70} has a mere five prime divisors. More delicate questions concern the distribution of f . If S is a set of positive integers, let $\mathbf{P}_N(S)$ be the fraction among the integers $1, 2, \dots, N$ that lie in S :

$$(13) \quad \mathbf{P}_N(S) = \frac{1}{N} \times \# [n: 1 \leq n \leq N \text{ and } n \in S].$$

The problem is to get information about quantities like $\mathbf{P}_N[n: a \leq f(n) \leq b]$.

Now (13) can be viewed as a probability: We draw an integer at random from the range $1 \leq n \leq N$, and $\mathbf{P}_N(S)$ is the probability that it will lie in S . That $\mathbf{P}_N[n: a \leq f(n) \leq b]$ can be viewed as a probability does not by itself ensure (this may be difficult to credit) that probability theory will help in the evaluation. It does in fact help because the notion of independence can be brought to bear. If $\delta_p(n)$ is 1 or 0 according as the prime p divides n or not, then $f(n) = \sum_p \delta_p(n)$. We can understand the distribution of $f(n)$ if we understand the joint behavior of the $\delta_p(n)$ as random quantities.

The number of multiples of p up to N is the integral part $[N/p]$ of N/p . The probability that $\delta_p(n) = 1$, or that $p \mid n$, is thus

$$(14) \quad \mathbf{P}_N[n: p \mid n] = \frac{1}{N} \left[\frac{N}{p} \right] \approx \frac{1}{p}.$$

The approximation here is good for large N : since $[N/p]$ differs from N/p by less than 1, the error in (14) is less than $1/N$. The formula (14) reflects the fact that p divides every p th integer, and it in no way requires that p be prime.

The fundamental theorem of arithmetic implies that, if integers a and b are relatively prime (share no prime factors), then they individually divide n if and only if their product ab divides n . This fact is well illustrated by the use Turing is said to have made of it. The sprocket wheel of his bicycle had a faulty tooth and the chain a faulty link, and unless he was pedalling very fast when the faulty parts meshed, the chain would fall off. So he counted the number, say a , of teeth on the wheel and the

number, say b , of links on the chain and found, not to his surprise, that a and b were relatively prime. Between successive meetings of the bad tooth and link the sprocket wheel would in consequence go through b cycles, as the chain went through a cycles. Turing is said to have pedalled along counting, on every b th cycle of the sprocket wheel giving the burst of speed necessary to carry him past the danger point.

As a special case of this fact, distinct primes p and q individually divide n if and only if pq does. By this and by (14) with pq in place of p ,

$$\mathbf{P}_N[n: p | n \text{ and } q | n] = \mathbf{P}_N[n: pq | n] = \frac{1}{N} \left[\frac{N}{pq} \right] \approx \frac{1}{pq} = \frac{1}{p} \cdot \frac{1}{q}.$$

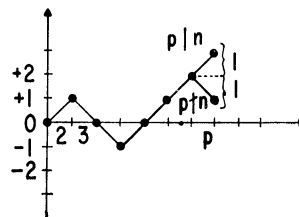
Since by (14) the factors $1/p$ and $1/q$ respectively approximate $\mathbf{P}_N[n: p | n]$ and $\mathbf{P}_N[n: q | n]$ if N is large, we arrive at

$$(15) \quad \mathbf{P}_N[n: p | n \text{ and } q | n] \approx \mathbf{P}_N[n: p | n] \mathbf{P}_N[n: q | n].$$

Thus the events $[n: p | n]$ and $[n: q | n]$ approximately satisfy the definition of independence if n is random, $1 \leq n \leq N$, with N large. There is an extension of (15) from two primes to three or more.

We can use this fact to construct a kind of random walk path containing information about the prime factorization of n and in particular about $f(n)$. We draw an integer n at random from among $1, 2, \dots, N$. On a vertical axis with the integer points marked off on it, we start at 0 and go up one unit if $2 | n$ and down one unit if $2 \nmid n$. From our new position ($+1$ or -1), we go up one unit if $3 | n$ and down one unit if $3 \nmid n$. We proceed in this way, examining each prime in succession. Figure 10 describes this factorization random walk in the same way that Figure 8 describes the coin-tossing random walk. Each number on the time axis is the prime corresponding to that step in the random walk. We consider later how long to continue the walk.

FIGURE 10



Since n is random, this path is random. But since the randomness is all in the drawing of n before the walk starts, the factorization random walk may seem less random than the coin-tossing random walk. This is an illusion. We may imagine tossing the coin T times in advance of the walk, recording the sequence of heads and tails, and only then performing the corresponding walk. Since we would see its whole history on record before setting out, the walk would be very dull. So imagine a friend who tosses the coin T times and records the results in advance of the journey, and imagine that, rather than show us the record all at once, he instead reveals the

outcomes to us one by one as we execute the walk. This restores the suspense. For the factorization random walk, we can imagine a friend who draws n at random, $1 \leq n \leq N$, factors n into primes, and at each step of the walk reveals to us whether or not the corresponding p divides n .

The increment of the random path in Figure 10 over an interval depends on how many in the corresponding set of primes divide n . Increments over disjoint intervals depend on disjoint sets of primes and hence by (15)—or by (15) together with its extension to three or more primes—the increments will be approximately independent if N is large. Unlike Brownian motion, however, the factorization random walk has a strong downward drift. By (14), the chance of going downward at the step corresponding to p is about $1 - 1/p$, which is almost 1 for large p . The remedy is to move up a distance $1 - 1/p$ if $p \mid n$ and to move down only a distance $1/p$ if $p \nmid n$. The expected distance moved is now

$$(1 - p^{-1})\mathbf{P}_N[n: p \mid n] + (-p^{-1})\mathbf{P}_N[n: p \nmid n],$$

which by (14) is approximately

$$\left(1 - \frac{1}{p}\right) \frac{1}{p} + \left(-\frac{1}{p}\right) \left(1 - \frac{1}{p}\right) = 0.$$

This corresponds with (8), an equation which shows that the coin-tossing random walk has no drift.

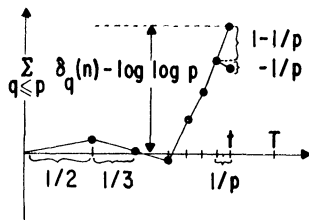


FIGURE 11

Since the mean distance moved at the step corresponding to p is approximately 0, the variance is approximately $(1 - p^{-1})^2 \mathbf{P}_N[n: p \mid n] + (-p^{-1})^2 \mathbf{P}_N[n: p \nmid n]$, which by (14) is in turn approximately

$$\left(1 - \frac{1}{p}\right)^2 \frac{1}{p} + \left(-\frac{1}{p}\right)^2 \left(1 - \frac{1}{p}\right) = \frac{1}{p} \left(1 - \frac{1}{p}\right) \approx \frac{1}{p}.$$

The distance moved thus tends to be very small for large p , in contrast with the coin-tossing random walk, which by (9) proceeds with vigor ever undiminished. The remedy this time is to spend only an amount of time $1/p$ executing the step corresponding to p . With these two modifications, the path is as in Figure 11.

To recapitulate, the time interval corresponding to the prime p has length $1/p$. Over this interval, the path rises an amount $\delta_p(n) - 1/p$; that is, it rises $1 - 1/p$ if $p \mid n$ (the probability of this is approximately $1/p$) and it rises $0 - 1/p$ if $p \nmid n$ (the probability of this is approximately $1 - 1/p$).

The point t in Figure 11 (the right endpoint of the interval corresponding to p) is $\sum_{q \leq p} 1/q$ (summation over primes q not exceeding p). The distance moved in the step corresponding to q has variance about $1/q$, and hence by the approximate independence of the steps ((15) again), the variance of the position at this time t is approximately $\sum_{q \leq p} 1/q$, or t itself. The above adjustment of the factorization random walk has thus not only eliminated the drift, it has so adjusted the time scale that the variances are about what they are for the coin-tossing random walk and for Brownian motion.

It can be shown that

$$(16) \quad \sum_{q \leq u} \frac{1}{q} \approx \log \log u$$

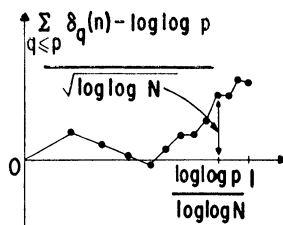
for large u (the two expressions go to infinity with u and their difference remains bounded; see [5, p. 351]). That the sum in (16), instead of increasing in some erratic fashion, is asymptotic to a standard function like $\log \log u$ is inessential to what follows; but the formulas become simpler (and remain valid) if at each occurrence of the sum we substitute the right side of (16).

Thus the t in Figure 11 is essentially $\log \log p$ and the height $\sum_{q \leq p} (\delta_q(n) - 1/q)$ of the curve over t is approximately

$$(17) \quad \sum_{q \leq p} \delta_q(n) - \log \log p.$$

Now n has $\sum_{q \leq p} \delta_q(n)$ prime divisors that do not exceed p , and we normalize this quantity by subtracting away the value $\log \log p$ it has for a "typical" n . (If n is random, $1 \leq n \leq N$, then $\sum_{q \leq p} \delta_q(n)$ has by (14) and (16) a mean of about $\log \log p$; this is where (12) comes from.) The factorization random walk is a record of these differences (17). We continue the walk until each $p \leq N$ has been dealt with, and the corresponding point on the time axis is $T = \sum_{p \leq N} 1/p \approx \log \log N$.

FIGURE 12



The random path now resembles a coin-tossing path in that the increments are almost independent for large N , there is essentially no drift, and the variances are about right. As in the coin-tossing case, rescaling will lead in the limit ($N \rightarrow \infty$) to Brownian motion. To send T to the point 1, we contract the horizontal scale by a factor $T = \log \log N$, and, again as in the coin-tossing case and for the same reasons, we contract the vertical scale by the square root of this, applying the transformation

(3). The point t in Figure 11 goes to $\log \log p / \log \log N$, and the path is that shown in Figure 12.

Since the path depends on n and N , denote it $\text{path}_N(n)$. Since n is random ($1 \leq n \leq N$), so is the path, and the chance that it lies in a given subset A of $C_0[0, 1]$ is $\mathbf{P}_N[n: \text{path}_N(n) \in A]$. The theorem linking primes with Brownian motion is this: If A is a subset (Borel subset) of $C_0[0, 1]$ satisfying $P_1(\partial A) = 0$, then

$$(18) \quad \mathbf{P}_N[n: \text{path}_N(n) \in A] \rightarrow P_1(A) \quad (N \rightarrow \infty),$$

where $P_1(A)$ is Wiener measure. The proof of (18) uses a combination of probability theory, functional analysis, and number theory. The theorem is given implicitly in [8, p. 122], explicitly in a manuscript version of [1] and in a much more general form in [9]. (For general discussions of probability methods in number theory, see [6], [8] and the author's 1973 Wald lectures, to appear in the *Annals of Probability*.)

From Figure 12, a plot of the differences (17) normalized to

$$(19) \quad (\sum_{q \leq p} \delta_q(n) - \log \log p) / \sqrt{\log \log N},$$

we can read off arithmetic properties of n , and therefore (18) yields arithmetic limit theorems. Consider the three sets A to which we applied the analogous result (10). The height of the curve in Figure 12 over the time point 1 is (19) with N in place of p ; it is the number $f(n)$ of prime factors of n , normalized to

$$(f(n) - \log \log N) / \sqrt{\log \log N}.$$

The greater this is, the more highly composite n is, and the smaller it is, the more "prime-like" n is. With $A = [x: \alpha \leq x(1) \leq \beta]$, it follows by (18) and (5) that

$$(20) \quad \mathbf{P}_N \left[n: \alpha \leq \frac{f(n) - \log \log N}{\sqrt{\log \log N}} \leq \beta \right] \rightarrow \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-u^2/2} du.$$

This is the Erdős-Kac central limit theorem for f . (For an elementary direct proof or (20), see [2].)

For $-\alpha = \beta = .9$, the limit in (20) is about .6, and if $N = 10^{70}$, so that $\log \log N \approx 5$, the double inequality in (20) is approximately the same as $-.9 \leq (f(n) - 5) / \sqrt{5} \leq .9$, which in turn is approximately the same as $3 \leq f(n) \leq 7$. Thus something like 60% of the integers under 10^{70} have from 3 to 7 prime divisors.

The larger (17) is, the more highly composite n appears to be at that point in the factorization; that is, (17) measures the apparent compositeness of n when it has been tested for divisibility only by primes up to p . The maximum apparent compositeness is measured by

$$(21) \quad \max_{p \leq N} \left(\sum_{q \leq p} \delta_q(n) - \log \log p \right);$$

since this is $\sqrt{\log \log N}$ times the maximum height of the curve in Figure 12, an application of (18) to the set in (6) gives its approximate distribution. The right side of (6) being about .1 if $\alpha = 1.7$, for about 10% of the integers under 10^{70} does (21) exceed $1.7 \times \sqrt{5} \approx 3.8$.

Let us say that n is *excessive at p* if

$$(22) \quad \sum_{q \leq p} \delta_q(n) > \log \log p;$$

this holds if, with respect to divisibility by primes up to p , n is “more composite,” or “less prime-like,” than the average integer. And (22) holds exactly when the corresponding point on the curve in Figure 12 is above the axis. The polygonal segment corresponding to p has length $p^{-1}/\log \log N$ when projected on the horizontal axis, and so the amount of time the curve spends above 0 is essentially

$$(23) \quad \frac{1}{\log \log N} \sum \left[\frac{1}{p} : p \leq N \text{ and } \sum_{q \leq p} \delta_q(n) > \log \log p \right],$$

the sum extending over those p at which n is excessive.

If we test n for divisibility by the primes in succession, spending an amount $1/p$ of time on p ($p \leq N$), (23) is the fraction of time we are dealing with a p at which n is excessive. From an application of (18) to the set in (7) it follows that for large N the distribution of (23) approximately follows the density curve in Figure 5. For about 20% of the integers under N the quantity (23) exceeds .9, for about 20% it is less than .1, and for only about 6% does it lie between .45 and .55.

Prime factors exhibit in this respect the same strange behavior coins do. In a way they are even more strange. A quantity perhaps more natural to consider than (23) is

$$(24) \quad \frac{1}{\pi(N)} \times \#_*[p: p \leq N \text{ and } \sum_{q \leq p} \delta_q(n) > \log \log p],$$

the *number* of p for which n is excessive at p , normalized by division by $\pi(N)$, the total number of primes involved. For N large, of the break points in the polygon in Figure 12 the great majority are very near 1, which has the result that in the limit the distribution of (24) consists of a mass of $\frac{1}{2}$ at 0 and a mass of $\frac{1}{2}$ at 1: If $\varepsilon > 0$ and N exceeds some N_ε , then (24) is less than ε with a probability lying in the range $\frac{1}{2} - \varepsilon$ and $\frac{1}{2} + \varepsilon$ and is greater than $1 - \varepsilon$ with a probability lying in the same range. Thus practically all integers are excessive either at practically all primes or at practically none.

The 1972 Rouse Ball Lecture, given while the author was a Guggenheim Fellow, visiting Peterhouse and the Statistical Laboratory of the University of Cambridge. It appeared in somewhat different form in *Eureka*, the Journal of the Archimedeans, the Cambridge University Mathematical Society.

References

1. Patrick Billingsley, *Convergence of Probability Measures*, Wiley, New York, 1968.
2. ———, On the central limit theorem for the prime divisor function, this MONTHLY, 76 (1969) 132–139.
3. William Feller, *An Introduction to Probability Theory and Its Applications*, vol. I, 3rd ed., Wiley, New York, 1968.
4. David Freedman, *Brownian Motion and Diffusion*, Holden-Day, San Francisco, 1971.
5. G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 4th ed., Clarendon Press, Oxford, 1960.
6. M. Kac, *Statistical Independence in Probability, Analysis and Number Theory*, Carus Math. Monogr. 12. MAA, Wiley, New York, 1959.
7. Samuel Karlin, *A First Course in Stochastic Processes*, Academic Press, New York, 1966.
8. J. Kubilius, *Probabilistic Methods in the Theory of Numbers*, 2nd ed. (1962). Vilna: Gosudarstv. Izdat. Politich. i Naučn. Lit. Litovsk. SSR. (English translation 1964. Amer. Math. Soc. Transl. of Math. Monographs, Volume 11.)
9. Walter Philipp, Arithmetic functions and Brownian motion, Proc. Symp. Pure Math., vol. 24, AMS, 1973.

CORRECTION TO “UNIQUE FACTORIZATION DOMAINS”

P. M. COHN, Bedford College, University of London

The statement “Any Noetherian UFD is a Dedekind domain” (this MONTHLY, 80 (1973) 1–18) should be omitted.

The assertion is of course well known to be false; a correct statement would be: A Dedekind domain is a UFD if and only if it is a principal ideal domain.

I am indebted to Professor J. H. Hays for drawing my attention to this error.

CORRECTION TO “A HISTORY OF THE PRIME NUMBER THEOREM”

L. J. GOLDSTEIN, University of Maryland

In my paper, [this MONTHLY, 80 (June-July, 1973) 599–615] I asserted that the sieve of Eratosthenes was known to the ancient Greeks and, in fact, appeared in Euclid. It has been pointed out to me by Professor J. Albree that although the sieve was known since approximately the time of Euclid, it does not appear in the *Elements*. The author regrets the error.

MATHEMATICAL NOTES

EDITED BY ROBERT GILMER

Material for this Department should be sent to David Roselle, Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803.

COMPLEMENTS AND COMMENTS

ROBERT GILMER AND DAVID ROSELLE

This annual article is designed to provide our readers with an outlet for remarks on papers that have appeared in the notes sections of the MONTHLY. Because of the nature of the material we seek to publish, it is inherent that there will be some duplication of results already in the literature; such duplication may not be undesirable, depending upon accessibility, presentation, etc. Under any circumstances, we are happy to have readers point out sources that are pertinent to articles we have published. During the past year, we have received the following information.

Calculus. The article *A unified proof of several basic theorems of real analysis* by Patrick Shanahan (this MONTHLY, 79,(1972) 895–8) has brought response from several readers. Leonard Gillman calls attention to the notes by L. R. Ford (this MONTHLY, 64 (1957) 106–8) and by W. L. Duren, Jr. (this MONTHLY 64 (1957) part II, 19–22). B. J. Pettis also credits the idea to Professor Ford and points out that the basic lemma appears in the calculus text by Brady and Mansfield. Professor Ray Glenn has pointed out that the ideas here also appear in *A creeping lemma* by R. M. F. Ross and G. T. Roberts (this MONTHLY, 75 (1968) 649–52). Daniel Mosenkis writes that Professor Jesse Douglas utilized this approach in notes distributed to his advanced calculus classes at C.C.N.Y. in the early part of the 1960's. An addendum to Shanahan's paper is scheduled to appear in the Classroom Notes section and it will examine some of these comments in more detail.

Richard Johnsonbaugh has noted that the technique used by J. P. Tull in *A discovery approach to e* (February, 1973, p. 193) has also been set forth by Flanders, Korfhage and Price in *A First Course in Calculus with Analytic Geometry*, pp. 204–5 and in *Calculus*, pp. 94–7. He has also pointed out that the approach taken by R. B. Darst in *Simple proofs of two estimates for e* (February, 1973, p. 194) is similar to that given by H. G. Eggleston in *Elements of Real Analysis*, Cambridge, 1962, pp. 24–5. An article on e by Johnsonbaugh is scheduled to appear in the Classroom Notes section in the near future.

Algebra. Several readers have commented on the proof of a characterization of injective modules contained in the article by Azmi Hanna (March, 1973, pp. 297–8). H. F. Kreimer points out that the characterization appears as exercise 5, page 371, of MacLane and Birkhoff's *Algebra*; James W. Brewer observes that the result in

question appears as Lemma 7, page 187, of *Rings and Fields* (Kaplansky) and as Theorem 3.18, page 48, of Joseph Rotman's *Notes on Homological Algebra* (Hanna's article supplies the same proof given in Rotman's book).

J. C. Ault has mentioned a couple of transcription errors in his article *Circle groups of nilpotent rings* (written jointly with J. C. Watters) in the January, 1973 MONTHLY (pp. 48–50). On page 49, $g+h$ should be defined to be $gh\{m(g, h)\}^{-1}$, and accordingly, on line 7 of page 50, $-g$ is $g^{-1}m(g, g^{-1})$ instead of $g^{-1}\{m(g, g^{-1})\}^{-1}$.

Walter Rudin has pointed out that his proof of the unique factorization theorem in $\mathbb{Z}[i]$ (this MONTHLY, 68 (1961) 907–8) is essentially the same as that given by M. F. Ruchte and R. W. Ryden in *A proof of uniqueness of factorization in the Gaussian integers* (January, 1973, pp. 58–9).

William H. Gustafson has supplied two references for the main result given by Claire Parkinson in *Ambivalence in alternating symmetric groups* (Feb., 1973, pp. 190–2). These are J. L. Berggren, *Finite groups in which every element is conjugate to its inverse*, Pacific J. Math., 28 (1969) 289–93, and A. Kerber, *Representations of permutation groups I*, Springer-Verlag, Lecture Notes, 240 (1971) 14. J. L. Brenner also writes that the main result of Parkinson's article follows from the known result that, in A_n , a permutation and its inverse are conjugate, except in the case that $n > 2$ and the parts of the cycle decomposition of the permutation are unequal and odd and an odd number of them are congruent to 3 modulo 4. Brenner does not supply us with a reference for this result, but we note that Kerber's volume contains a treatment.

The main result of the article on convex matrix functions by M. H. Moore (April, 1973, pp. 408–9) has a long history. It has previously appeared in each of the following articles: *A multivariate generalization of Tchebichev's inequality*, by P. Whittle, Quart. J. Math., (2) 9 (1958) 232–240; *A multivariate Tchebycheff inequality*, by I. Olkin and J. W. Pratt, Ann Math. Stat., 29 (1958) 226–234; *On the bias of functions of the characteristic roots of a random matrix*, by T. Cacoullus and I. Olkin, Biometrika 52 (1965) 87–94; *A note on the expected value of the inverse of a matrix*, by T. Groves and T. Rothenburg, Biometrika, 56 (1969) 690; *On the expectation of the inverse of a matrix*, by V. K. Srivastava, Sankhya (Series A) 32 (1970) 336. This information was transmitted by M. D. Perlman and I. Olkin.

In his historical article on the invariance of parity of permutations (August–September, 1972, pp. 776–9), T. L. Bartlow makes the following statement. “When a product of cycles is multiplied by a transposition, the number of cycles of the product is either one greater or one less.” He “proves” the statement by simply writing down a formula. J. L. Brenner objects, on logical grounds, that Bartlow should establish his formula by induction, stating clearly what assumptions concerning permutations are being made.

Analysis. In their paper *A weak parallelogram law for l_p* (November, 1972, pp. 1012–1015) W. L. Bynum and J. H. Drew attribute the inequality

$(\|x\| + \|y\|)^p + \|\|x\| - \|y\|\|^p \leq \|x + y\|^p + \|x - y\|^p$; $1 < p \leq 2$, $x, y \in l_p$ to O. Hanner in his article in *Ark. Mat.*, 3 (1956) 239–244. But R. Askey points out that Hanner indicates the preceding inequality is due to A. Beurling; Hanner repeats Beurling's proof in his paper.

Jack Vilms observes that the orthogonality condition of R. K. Williams in his March, 1973 note on conformality of mappings is quite natural, when placed in the proper perspective. A sketch of Vilms' approach is the following. Interpret conformality of a mapping $f: \mathbb{R} \rightarrow \mathbb{R}$ in terms of conformality of the Jacobian matrix of f . Then taking, in Williams' notation,

$$L_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, L_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, L_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, L_4 = \begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

Vilms obtains Williams' criterion.

The December, 1972 MONTHLY contained an article by M. A. Golberg on the derivative of a determinant. P. Lancaster and J. Rokne indicate that the result Golberg labels 'Main Theorem' has a previous history. In his article *The evaluation of determinants by the method of variation of parameters*, Soviet Math. Doklady I (1960) 316–9, D. F. Davidenko attributes the result to A. Halanai; a proof of the theorem that may be more accessible appears as Appendix II of Lancaster's paper *Algorithms for lambda-matrices*, Numerische Math., 6 (1964) 388–394. Golberg's Main Theorem also appears as exercise 7, page 195, of Lancaster's book *Theory of Matrices*; his corollary 2 appears as exercise 8 on the same page.

General. J. G. Mauldon calls attention to an interesting source that is likely to be inaccessible to most MONTHLY readers — a high school trigonometry textbook, entitled *Advanced Trigonometry*, by C. V. Durell and A. Robson. Mauldon states that the elegant style of the book, which was published in London in 1930, was an inspiration to many English school children. The book is mentioned here because of its relation to two articles that appeared in the April, 1973 MONTHLY, one by I. Papadimitriou and the other by T. M. Apostol. Papadimitriou gives a simple proof of the formula $\sum_{k=1}^{\infty} k^{-2} = \pi^2/6$, and Apostol generalizes to the series $\sum_{k=1}^{\infty} k^{-2n}$. In Example 7, page 209, Durell and Robson present an elementary proof of the equality $\sum_{k=1}^{\infty} k^{-2} = \pi^2/6$; their methods are the same as those of Papadimitriou and Apostol. Mauldon indicates that subsequent exercises in Durell-Robson give other results of Apostol's paper.

'Topology. Solomon Leader and Ernest Michael have (independently) observed that Theorem 1 of the paper *A theorem on set inclusion in metric spaces*, by James Heinen and Albert Wilansky, in the January, 1973 MONTHLY carries over to a general topological space X that is not necessarily metrizable. Michael states that metrizability is used in the proof only in proving that D_2 is nonempty, and there the assumption is easily avoided.

A PROBLEM ON SERIES

G. J. O. JAMESON, Universität Innsbruck

For a sequence x of real or complex numbers, we denote by $x(n)$ the n th term of x . We prove the following result.

THEOREM. *If $\sum |x(n)|$ is convergent and $\sum_{n=1}^{\infty} x(kn) = 0$ for each positive integer k , then x is the zero sequence.*

Proof. Let p_n be the n th prime greater than 1, and let R_n be the set of positive integers that are not divisible by any of p_1, \dots, p_n . Note that the first two members of R_n are 1 and p_{n+1} . We show that

$$(1) \quad \sum_{j \in R_n} x(j) = 0$$

for each n . From this it follows that $|x(1)| \leq \sum_{p_{n+1}}^{\infty} |x(j)|$ for all n , and therefore that $x(1) = 0$. A similar argument shows that, for any k , $\sum_{j \in R_n} x(kj) = 0$, and hence that $x(k) = 0$.

For a bounded sequence y , write $\langle x, y \rangle = \sum x(n)y(n)$. Let a_k be the sequence with 1 in place nk ($n = 1, 2, \dots$) and 0 elsewhere. The hypothesis states that $\langle x, a_k \rangle = 0$ for each k .

Choose n , and for $1 \leq r \leq n$, let T_r be the set of products of r distinct primes chosen from p_1, \dots, p_n . Let

$$b = a_1 + \sum_{r=1}^n (-1)^r \sum_{k \in T_r} a_k.$$

We show that $b(j)$ is 1 for $j \in R_n$ and 0 for other j ; statement (1) then follows. If j is in R_n , then $a_k(j) = 0$ for all k in $\bigcup_{r=1}^n T_r$, so $b(j) = a_1(j) = 1$. Suppose now that j is not in R_n , and let q_1, \dots, q_m be the primes that divide j and are not greater than p_n . Let U_r be the set of products of r distinct primes chosen from q_1, \dots, q_m . Then U_r has $\binom{m}{r}$ members, and $a_k(j) = (-1)^r$ for each k in U_r . These, together with a_1 , are the only a_k that make a non-zero contribution to $b(j)$, which is therefore equal to

$$1 + \sum_{r=1}^m (-1)^r \binom{m}{r} = (1-1)^m = 0.$$

This completes the proof.

PROBLEM. Does the conclusion hold if we drop the assumption that $\sum |x(n)|$ is convergent?

Added in proof: This problem has been solved negatively by D. H. Fremlin.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics, Statistics, and Computing Science, The University of Calgary, Calgary 44, Alberta, Canada.

MONTHLY RESEARCH PROBLEMS, 1969–73

RICHARD K. GUY, The University of Calgary

This continues the article written with Victor Klee [1971, 1113], referred to hereafter as GK. References to papers appearing in this section of the Monthly are in brackets containing the year and first page; they are not listed at the end. Other references are in parentheses; a date or tbp (to be published) refers to the list at the end of the article; wrc (written communication) is used when there are no (known) publication plans.

The editor welcomes comments on the Research Problems section. It has been suggested that since the MONTHLY is not a research periodical, research problems are out of place. On the other hand, there is the view that it is important for students and teachers to realize that though they may not (yet) feel capable of mathematical research, there are plenty of unsolved problems that are well within their comprehension. Many amateurs are attracted to the subject, and many successful researchers first gained their confidence by examining intuitive problems in areas such as euclidean geometry, theory of numbers and, latterly, combinatorics and graph theory, where it is possible to understand questions (even to formulate them) and to obtain original results without a deep prior theoretical knowledge. Many years ago a friend telephoned me that he had got into an argument with his wife as to whether there was an infinity of primes. I find it significant that people without special mathematical training should have initiated such a discussion; moreover that I was able to outline Euclid's proof to the enquirer's satisfaction over the telephone.

A point that occasionally arouses controversy is that the Research Problems section does not publish *solutions*. Occasionally, when space permits, solutions do appear, usually in the Mathematical Notes section, but normally papers should be submitted to journals regularly devoted to research. It is hoped that the problems appearing in this section are of interest to a reasonably wide band of the broad spectrum of MONTHLY readers, many of whom are undergraduates, college teachers and others, by no means all active in research. When a problem is solved, the solution, especially the first one, may be complicated, or prolix, or sophisticated, or just dull, and unsuitable for publication in the MONTHLY. To take an extreme example, the four

color conjecture is suited to these pages on grounds of intelligibility, though not on others; the solution, when it appears, may be quite unsuited. However, I hope that those who solve or make partial progress with problems in this section will let me have offprints or preprints, or other comments and references, so that future updating articles can be as complete and accurate as possible.

Carmichael's conjecture, concerning the non-uniqueness of solutions of the equation $\phi(x) = n$, where ϕ is Euler's totient function, was discussed by Klee [1969, 288]. Grosswald (1973) has proved that the conjecture holds, except possibly for integers n divisible by 32; this is a slight strengthening of a result of Donnelly (tbp), who also shows that if x_0 is the least integer x for which $\phi(x) = n$ has a unique solution then $n = \phi(x_0) \equiv 0 \pmod{2^{14}}$.

Larman (1971) answered a question of Klee [1969, 678], by proving that all convex borel sets in E^3 can be generated in a borelian manner within the realm of convexity. He now writes that the same answer has recently been proved in E^d , $d \geq 2$, by David Preiss of Prague.

Kronk [1969, 809] asked if there exists a hypotraceable graph. J. D. Horton (tbp) has constructed one from five copies of the Petersen graph (Figure 1).

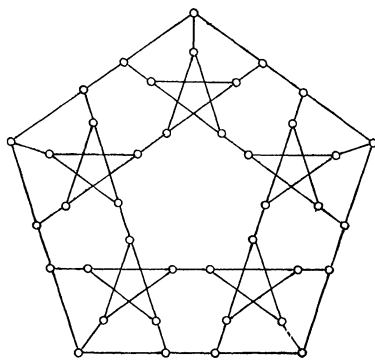


FIG. 1

Kronk [1969, 1045] discussed the conjecture of Nash-Williams and Plummer that the square of every non-separable graph is hamiltonian. A number of proofs have been combined in a paper by G. Chartrand, A. M. Hobbs, H. A. Jung, S. F. Kapoor and C. St. J. A. Nash-Williams (tbp), whose main results are that the square of every 2-connected graph is hamilton-connected and that the square of every 2-connected graph of order at least 4 is 1-hamiltonian. This work stems from that of Fleischner (tbp); an exact reference can now be given to Fleischner and Kronk (1972). Chartrand writes that Zaks (1972) has shown that if G is 2-connected, G^2 need not be 2-hamiltonian, and that Hobbs (1973) has shown that G^2 is vertex pancyclic.

I am indebted to Golomb, Rosa, Sheppard, Stanley and Stanton for helpful correspondence concerning Duke's [1969, 1128] paper on Ringel's (1963) tree-packing problem. Sheppard (wrc) has given a counterexample to the sufficiency condition mentioned in GK as having been suggested by Haggard and McWha. We referred in GK to Kotzig and Rosa (1970), where Rosa (1966) would have been more appropriate. The former define a **magic valuation** as a labelling of the vertices and edges of a graph so that the total of the labels on each edge and its incident vertices is (the same 'magic') constant. What is needed here was called by Golomb (1972) a **graceful numbering**; a labelling of the vertices with non-negative integers so that the absolute differences of the ends of the e edges comprise the first e positive integers. There is no loss of generality in assuming that one of the vertices is labelled zero; such a graceful numbering was called a **β -valuation** by Rosa (1966). He used the name **α -valuation** if, in addition, there was a number lying between (not strictly) the labels of the two vertices incident with any edge; i.e., only bipartite graphs can have α -valuations. Trees are bipartite, but not all trees have α -valuations; Rosa gives the example of Figure 2. However, he found β -valuations (graceful numberings) for all

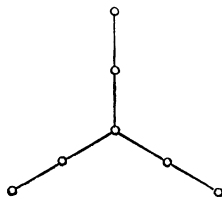


FIG. 2.

trees with 16 or fewer vertices. Kotzig (1972b) proved the following important result: given a tree, form an infinite family of trees by replacing each edge in turn by a path of arbitrary length; then such a family contains only a finite number of trees with no α -valuation. In a recent paper Stanton and Zarnke (tbp) show that if S and T are gracefully numbered trees, then the balanced trees formed from S and T can also be so labelled; a **balanced tree** of type I (resp. II) being formed by attaching a copy of T to every vertex (resp. with one exception) of S . Kotzig (1972a) discusses the relationship between α -valuations and magic valuations, but this paper is more concerned with the latter. A further use of **magic labelling** is that of Stanley (tbp) who labels the edges and requires the edge-sums at each vertex to be the same. This last is also the usage of Stewart (1966) whose interest derived from a problem of Sedláček (1963). Murty [1971, 1000] also used the word 'magic' on the encouragement of the present writer; we do not know of any progress with his problem.

The asymptotic result on the problem discussed by Klee [1970, 63] on the longest d -dimensional snake has now appeared as Wyner (1971). Adelson, Alter, and Curtz (1973) have found new snakes of lengths 48, 86 and 128 in 7, 8 and 9 dimensions; the last two are longer than any previously known. Klee reports that Preparata and Nievergelt (tbp) have obtained new results for this problem.

The paper of Subbarao, Cook, Newberry and Weber (1972) on unitary perfect numbers [1970, 389] has now appeared, as has that of Usiskin and Wayment (1972) on partitioning a triangle into 5 triangles similar to it [1970, 867]. The more general problem [1971, 1118] of partitioning a triangle into 5 similar triangles has been examined by R. B. Killgrove (wrc) and in exhaustive detail by Don Coppersmith (wrc). There are evidently 10 essentially different configurations: Figure 3 is any isosceles triangle dissected into right triangles; Figures 4 and 5 are equilateral triangles, one dissected into right triangles, the other into triangles containing an angle of 120° ; Figure 6 is a $90^\circ, 60^\circ, 30^\circ$ triangle dissected into $120^\circ, 30^\circ, 30^\circ$ ones. The other 6 cases are dissections of various scalene triangles of prescribed shape, 2 of them containing an angle of 60° ; in these 2 the constituent triangles each contain an angle of 120° .

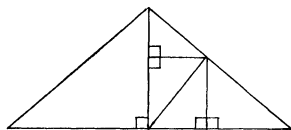


FIG. 3.

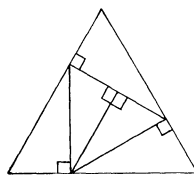


FIG. 4.

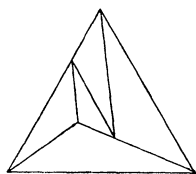


FIG. 5.

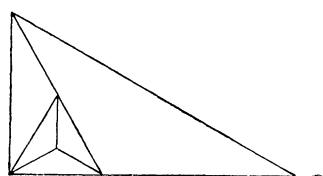


FIG. 6

The solution of Fejes Tóth's [1970, 869] illumination problem has been completed by Bruce Henry (1973).

Wills writes that the paper of Bokowski and Odlyzko (1973) is relevant to his problem [1971, 47] on lattice points and volume/area ratio of convex bodies, and that Bokowski, Hadwiger and Wills (1972) have solved the related problem of finding a lower bound for the number of lattice points, $G(K)$, of a convex body. There are two conjectures concerning the problems of finding an upper bound for $G(K)$ and for $G(\dot{K})$, the number of lattice points on the boundary. The first,

$$(?) \quad G(K) \leq \sum_{i=0}^n \binom{n}{i} \frac{W_i(K)}{w_i},$$

where $W_i(K)$ is Minkowski's "*Quermassintegrale*" and w_i is the volume of the i -dimensional unit sphere, is mentioned in Wills (1973). The second, for $G(\dot{K})$, is due to Hadwiger and Wills; the sum is taken over only odd values of i and then

doubled. Partial results have been obtained by Wills in collaboration with Bokowski (1973), McMullen (1973) and Hadwiger (1973).

Rosenfeld's problem [1971, 49] to find the number of graphs on n vertices with k cliques has been solved for $k = 3$ by P. McMullen (wrc); the number may be written

$$[(n+3)(6n^4 - 18n^3 + 34n^2 - 62n + 165 + 45(-1)^n)/1440],$$

where brackets denote greatest integer not greater than.

Doran's (1972) solution to his problem [1971, 178]: if A is a symmetric $*$ -algebra without identity, is the $*$ -algebra obtained from A by adjoining an identity symmetric; has now appeared.

Lind writes that Cahen informs him that his problem [1971, 179] on which polynomials map the algebraic integers into themselves, goes back to Ostrowski (1919) and Pólya (1919). Cahen has sent a description of his work (1972, 1973), that of Chabert (1971, 1972, 1973), of Brizolis (1973, tbp), of Gunji and McQuillan (1969, 1970) and of McClure (1971).

Hering's (1973) paper solving his problem [1971, 275] on inequalities has now appeared, and Singmaster's (1973) paper concerning his problem [1971, 385] on repetitions of integers as binomial coefficients is also appearing.

The downfall of Duke's [1971, 386] conjecture concerning the genus and Betti number of a graph is mentioned by Nordhaus (1972). His papers with Stewart and White (1971) and with Ringeisen, Stewart and White (1972) are related. Duke mentions two other papers of Ringeisen (1972) and states his belief that he can use the work of Martin Milgram and Peter Ungar, reported in GK [1971, 1120], to show that the best possible relation between the Betti number and genus of a graph is $\beta \geq \gamma(2 + c/\log \gamma)$ for some constant c .

Witsenhausen (tbp) has improved Rosenthal's bounds, quoted by Bolker [1971, 529] in discussing the zonoid problem.

Payne writes concerning his article [1971, 659] on linear transformations of a finite field that his paper (1972) has now appeared, and that he has solved (1971) the problem (not the general problem) mentioned there.

Chui (tbp) has a paper related to his problem [1971, 779] on fields, due to point masses, as does Newman (tbp).

Herda's conjecture [1971, 888] that a circle maximizes the minimum pseudo-diameter has been confirmed by Ault (1974), Batten (wrc), Chakerian (1974), Davies (wrc), Fink (wrc), Goodey (1972), Johnson (wrc), Lipskie (wrc), Short (wrc), Wenté (wrc) and Witsenhausen (1972); see Herda's (1974) article for details. A related problem was considered by Besicovitch (1961), Danzer (1963), Koenen (1971) and Nash-Williams (1972).

Smith writes that the two conjectures of his paper with Kumin [1972, 157] have been answered affirmatively by Joan P. Hutchinson (tbp), Frank Owens and Louis H. Rowen (tbp).

Papers related to Peterson's question [1972, 505]: do self-intersections characterize curves of constant width, are those of Goodey (tbp) and Peterson himself (tbp).

H. Kharaghani (wrc) has studied the Hadamard maximum determinant problem discussed by Brenner and Cummings [1972, 626]: the reference to Schmidt (1970) has wrong page numbers and a reference to Hall and Ryser (1951) might have been more appropriate than the one to Hall (1956). The bound given by Popoviciu (1937) is not, as stated, sharper than that of Barba (1933). As a start to the related problem that they mentioned, in which matrix entries are restricted to a sector $|\theta| \leq \theta_0 \leq \pi$ of the unit circle, Brenner (wrc) and Cummings show that for $n = 2$ the maximum modulus of the determinant is $\max\{2, 2 \sin 2\theta_0\}$.

Doran writes to point out that the problem in his paper [1972, 762], does there exist more than one Banach *-algebra with discontinuous involution, has some trivial solutions: e.g., adjoin an identity twice to Bonsall's example and take finite direct sums, or take the tensor product with a finite-dimensional algebra. These two constructions can be carried out before or after adjoining an identity, but are of more interest in the latter case.

The footballers of Croam, discussed by Biggs [1972, 1020] do not need to play on Sunday. The conjecture that O_k is never edge- k -colorable is false. Meredith and Lloyd (1972, tbp) have shown that k colors suffice for the edges of O_k when $k = 5$ and 6. The case $k = 5$ was also settled by G. Szekeres. In fact Meredith and Lloyd show that O_4 is the union of 2 hamilton circuits, O_5 the union of 2 hamilton circuits and a 1-factor and O_6 the union of 3 hamilton circuits, so that Biggs is now tempted towards the opposite conjecture, that for $k > 3$, O_k contains $\lfloor \frac{1}{2}k \rfloor$ edge-disjoint hamilton circuits.

There is a misprint in the paper of Erdős and Guy [1973, 52]; the ratio of the crossing number of the complete graph to $\binom{n}{4}$ (not n^4) tends to a limit between $3/10$ and $3/8$. D. Singer (wrc) has shown that the rectilinear crossing number of the complete graph on 10 vertices satisfies $\bar{v}(K_{10}) \leq 62$, and he and H. F. Jensen (wrc) have independently improved Jensen's (1971) upper bound. From Jensen's work it is possible to deduce that

$$\bar{v}(K_n) \leq [(n-1)(n-3)^2(5n-4)/312]$$

and each notes that the limit corresponding to that mentioned above is at most $5/13$. In view of Singer's discovery, it is now more plausible to conjecture that this upper bound can be further improved. We reported that the number of non-isomorphic optimal drawings of K_9 is 'about 200'. This was based on the fact that there are 181 drawings in which the **responsibility** of (total number of crossings on edges incident with) at least one vertex is 18, and the belief that there were few drawings with all responsibilities smaller. A. Uytterhoeven (wrc) of Baal, Belgium, has made an extensive search for such drawings. He confirms the number 181, but has found no fewer than 230 drawings with responsibilities less than 18, 4 of them with all

responsibilities 16. He does not claim completeness, but says that the list of 411 '*est à peu près complète*'.

By way of repeating the plea for help with keeping the section up-to-date by readers' comments, references, preprints and offprints, I conclude by saying that most of this article is owed to a large number of helpful correspondents; not only those mentioned by name or for whom space does not permit a mention, but also editors of journals and referees, for whom confidentiality demands that their unrewarding work goes unacknowledged but not unappreciated.

References

- L. E. Adelson, R. Alter and T. B. Curtz (tbp), Long snakes and a characterization of maximal snakes on the d -cube, *Congressus Numeratum VIII Proc. 4th S. E. Conf. on Combinatorics, Graph Theory and Computing*, Boca Raton, 1973.
- , Computation of d -dimensional snakes, *Congressus Numeratum VIII Proc. 4th S. E. Conf. on Combinatorics, Graph Theory and Computing*, Boca Raton, 1973.
- R. Ault, Metric characterization of circles, this MONTHLY, 81 (1974) to appear.
- G. Barba, Intorno al teorema di Hadamard sui determinanti a valore massimo, *Giorn. Mat. Battaglini*, 71 (1933) 70–86.
- A. S. Besicovitch, A problem on a circle, *J. London Math. Soc.*, 36 (1961) 241–244.
- J. Bokowski, H. Hadwiger and J. M. Wills, Eine Ungleichung zwischen Volumen, Oberfläche und Gitterpunktzahl konvexer Körper im n -dimensionalen euklidischen Raum, *Math. Z.*, 127 (1972) 363–364.
- and A. M. Odlyzko, Lattice points and the volume/area ratio of convex bodies, *Geom. Dedicata*, 2 (1973).
- and J. M. Wills, Upper bounds for the number of lattice points of convex bodies, this MONTHLY, 81 (1974) to appear.
- D. Brizolis, Ideals of rings of integer valued polynomials, Ph. D. dissertation, U. C. L. A., 1973.
- (tbp) On the ratios of integer-valued polynomials over any algebraic field.
- P. J. Cahen, Polynômes à valeurs entières, *Canad. J. Math.*, 24 (1972) 747–754.
- , Polynômes à valeurs entières, Thèse, Paris, 1973.
- and J. L. Chabert, Coefficients et valeurs d'un polynôme, *Bull. Sci. Math.*, 95 (1971) 295–304.
- J. L. Chabert, Anneaux de polynômes à valeurs entières et anneaux de Fatou, *Bull. Soc. Math. France*, 99 (1972) 273–283.
- , Anneaux de polynômes à valeurs entières, *Colloq. d'alg. Rennes*, No. 8 (1972).
- , Anneaux de polynômes à valeurs entières et extensions de Fatou, Thèse, 1973.
- G. D. Chakerian, A characterization of curves of constant width, this MONTHLY, 81 (1974) to appear.
- G. Chartrand, A. M. Hobbs, H. A. Jung, S. F. Kapoor and C. St. J. A. Nash-Williams (tbp), The square of a block is hamiltonian connected, *J. Combinatorial Theory*.
- C. K. Chui (tbp), On approximation in the Bers spaces, *Proc. Amer. Math. Soc.*,
- L. Danzer, A characterization of the circle, *Proc. Sympos. Pure Math.*, VII, Convexity, Amer. Math. Soc., 1963, 99–100.
- H. Donnelly (tbp), On a problem concerning Euler's phi-function, this MONTHLY, 80 (1973) 1029–1031.
- R. S. Doran, A generalization of a theorem of Civin and Yood on Banach*-algebras, *Bull. London Math. Soc.*, 4 (1972) 25–26.

- H. Fleischner (tbp), On spanning subgraphs of a connected bridgeless graph and their application to DT-graphs, *J. Combinatorial Theory*, 16B (1974).
- , The square of every two-connected graph is Hamiltonian, *ibid.*
- H. Fleischner and H. B. Kronk, Hamiltonsche Linien in Quadrat brückenloser Graphen mit Artikulationen, *Monatsh. Math.*, 76 (1972) 112–117.
- S. W. Golomb, How to number a graph, in R. C. Read (ed.), *Graph Theory and Computing*, Academic Press, 1972, 23–37.
- P. R. Goodey, A characterization of circles, *Bull. London Math. Soc.*, 4 (1972) 199–201.
- (tbp) Intersections of circles and curves of constant width.
- E. Grosswald, Contribution to the theory of Euler's function $\phi(x)$, *Bull. Amer. Math. Soc.*, 79 (1973) 337–341.
- H. Gunji and D. L. McQuillan, On polynomials with integer coefficients, *J. Number Theory*, 1 (1969) 486–493.
- , ———, On a class of ideals in an algebraic number field, *J. Number Theory*, 2 (1970) 207–221.
- H. Hadwiger and J. M. Wills, Über Eikörper und Gitterpunkte in gewöhnlichen Raum, *Geom. Dedicata*, 2 (1973).
- M. Hall, A survey of difference sets, *Proc. Amer. Math. Soc.*, 7 (1956) 975–986.
- M. Hall and H. J. Ryser, Cyclic incidence matrices, *Canad. J. Math.*, 3 (1951) 495–502.
- B. R. Henry, Solution of Fejes Tóth's illumination problem, this MONTHLY, 80 (1973) 409–410.
- H. Herda, A characterization of circles and other closed curves, this MONTHLY, 81 (1974) to appear.
- F. Hering, Eine Verallgemeinerung der Ungleichung vom arithmetischen und geometrischen Mittel, *Monatsh. Math.*, 77 (1973) 31–42.
- A. M. Hobbs, The square of a block is vertex pancyclic, *Graph Theory Newsletter*, W. Mich. Univ., 2 #5 (1973) 2.
- J. P. Hutchinson (tbp), Eulerian graphs and polynomial identities for sets of matrices; see also *Matrices satisfying $[A_1, A_2, \dots, A_n] = 0$* , *AMS Notices*, 19 (1972) A729.
- H. F. Jensen, An upper bound for the rectilinear crossing number of the complete graph, *J. Combinatorial Theory*, 10B (1971) 212–216.
- W. Koenen, Characterizing the circle, this MONTHLY, 78 (1971) 993–996.
- A. Kotzig, On vertex-valuations and magic valuations of certain bichromatic graphs, *Publications C. R. M.* #233, Montreal, Oct. 1972.
- , On certain vertex valuations of finite graphs, *Publications C. R. M.* #236, Montreal, Oct. 1972.
- and A. Rosa, Magic valuations of finite graphs, *Canad. Math. Bull.*, 13 (1970) 451–461.
- D. G. Larman, The convex borel sets in R^3 are convexly generated, *J. London Math. Soc.*, (2), 4 (1971) 5–14.
- C. R. McClure, Common divisors of values of polynomials, *J. Number Theory*, 3 (1971) 33–34.
- P. McMullen and J. M. Wills, Zur Gitterpunktanzahl auf dem Rand konvexer Körper, *Monatsh. Math.*, 77 (1973).
- G. H. J. Meredith and E. K. Lloyd, The hamiltonian graphs O_4 to O_7 , in D. J. A. Welsh (ed.), *Combinatorics*, I. M. A. 1972, 229–236.
- , (tbp), The footballers of Croam, *J. Combinatorial Theory*.
- C. St. J. A. Nash-Williams, Plane curves with many inscribed rectangles, *J. London Math. Soc.*, 5 (1972) 417–418.
- D. J. Newman (tbp), A lower bound for an area integral.
- E. A. Nordhaus, On the girth and genus of a graph, in *Graph Theory and Applications (Proc. Conf. W. Mich. U.)*, Springer, 1972, 207–214.

- , B. M. Stewart and A. T. White, On the maximum genus of a graph, *J. Combinatorial Theory*, 11B (1971) 258–267.
- , R. D. Ringeisen, B. M. Stewart and A. T. White, A Kuratowski-type theorem for the maximum genus of a graph, *J. Combinatorial Theory*, 12B (1972) 260–267.
- A. M. Ostrowski, Über ganzwertige Polynome in algebraischen Zahlkörpern, *J. Reine Angew. Math.*, 149 (1919) 117–124.
- S. E. Payne, A complete determination of translation ovoids in finite desarguan planes, *Atti Accad. Naz. Lincei, Rend. Cl. Sci. Fis. Mat. Natur.*, 51 (1971) 328–331.
- , Generalized quadrangles as amalgamations of projective planes, *J. Algebra*, 22 (1972) 120–136.
- B. B. Peterson (tbp), Intersection properties of curves of constant width, *Illinois J. Math.*
- G. Pólya, Über ganzwertige Polynome in algebraischen Zahlkörpern, *J. Reine Angew. Math.*, 149 (1919) 97–116.
- J. Popoviciu, Remarques sur le maximum d'un déterminant dont tous les éléments sont non négatifs, *Bul. Soc. Ști. Cluj*, 8 (1937) 572–582.
- R. D. Ringeisen, Determining all compact orientable 2-manifolds upon which $K_{m,n}$ has 2-cell embeddings, *J. Comb. Theory*, 12B (1972) 101–104.
- , Upper and lower embeddable graphs, in *Graph Theory and Applications*, Proc. Conf. W. Mich. Univ., 1972, 261–268.
- G. Ringel, Problem 25, *Theory of Graphs and its Applications*, Proc. Sympos. Smolenice 1963, Prague, 1964, 162.
- A. Rosa, On certain valuations of the vertices of a graph, *Theory of Graphs*, Proc. Internat. Sympos. Rome, 1966, Gordon & Breach, N. Y., 1967, 349–355.
- L. H. Rowen, On classical quotients of polynomial identity rings with involution, *Proc. Amer. Math. Soc.*, 40 (1973) 23–29; see also Standard identities for matrix rings with involution, *AMS Notices*, 20 (1973) A76.
- K. W. Schmidt, Lower bounds for maximal (0,1) determinants, *SIAM J. Appl. Math.*, 19 (1970) 440–442.
- J. Sedláček, Problem 27, *Theory of Graphs and its Applications*, Proc. Sympos. Smolenice 1963, Prague, 1964, 163–164.
- D. Singmaster, Repeated binomial coefficients and Fibonacci numbers, *Fibonacci Quart.*, 11 (1973).
- R. Stanley (tbp), Linear homogeneous diophantine equations and magic labellings of graphs, *Duke Math. J.*, 40 (1973).
- R. G. Stanton and C. R. Zarnke, (tbp), Labelling of balanced trees, *Congressus Numeratum VIII Proc. 4th S. E. Conf. on Combinatorics, Graph Theory and Computing*, Boca Raton, 1973.
- B. M. Stewart, Magic graphs, *Canad. J. Math.*, 18 (1966) 1031–1059.
- M. V. Subbarao, T. J. Cook, R. S. Newberry and J. M. Weber, On unitary perfect numbers, *Delta*, 3 (1972) 22–26.
- Z. Usiskin and S. G. Wayment, Partitioning a triangle into 5 triangles similar to it, *Math. Mag.*, 45 (1972) 37–42.
- J. M. Wills, Zur Gitterpunktanzahl konvexer Mengen, *Elem. Math.*, 28 (1973).
- H. S. Witsenhausen, On closed curves in Minkowski spaces, *Proc. Amer. Math. Soc.*, 35 (1972) 240–241.
- , (tbp), Metric inequalities and the zonoid problem, *Proc. Amer. Math. Soc.*,
- A. D. Wyner, Note on circuits of spread k in the n -cube, *I. E. E. E. Trans. Computers*, C20 (1971) 474.
- J. Zaks, *Graph Theory Newsletter*, W. Mich. Univ., 1 #4 (1972) 7.

CLASSROOM NOTES

EDITED BY ROBERT GILMER

Material for this Department should be sent to David Roselle, Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803.

THE MINIMAL POLYNOMIAL OF A LINEAR TRANSFORMATION

M. D. BURROW, Courant Institute of Mathematical Sciences

1. Introduction. It seems that none of the textbooks on linear algebra gives a direct proof of the fact that the minimal polynomial $m(x)$ of a linear transformation T is of degree less than or equal to the dimension of the vector space V on which T acts. The usual proof depends on the fact that $m(x)$ divides the characteristic polynomial $f(x)$, the degree of which is equal to the dimension of V , and for this one needs the Cayley-Hamilton theorem. In Theorem 1 we give a direct proof using mathematical induction on the dimension of V . The induction is brought into play by the use of quotient spaces and linear transformations induced on them by invariant subspaces. Theorem 2 shows that if $m(x)$ is a power of an irreducible polynomial $p(x)$ of degree r , then r divides n , where n is the dimension of the vector space V . This leads to an expression for the characteristic polynomial $f(x)$ in terms of the irreducible factors of $m(x)$ in the general case.

2. THEOREM 1. *Let V be a vector space of dimension n over a field F . Let $T: V \rightarrow V$ be a linear transformation. Then the minimal polynomial $m(x)$, that is the monic polynomial of minimal degree for which $m(T)V = 0$, is of degree less than or equal to n .*

Proof. Suppose that $\dim V = 1$. Then for any non-zero vector α in V we have $V = F\alpha$. It follows that $T\alpha = k\alpha$, where k is some element of F . Hence $(T - kI)\alpha = 0$ where I is the identity map on V . This shows that $m(x) = x - k$ is the minimal polynomial of T . Since $\deg m(x) = 1$ we see that the theorem is true for the case $n = 1$.

Now, to make an induction on the dimension we assume that the theorem is true for all spaces W of dimension less than n . Let $\dim V = n$, and suppose that α is a non-zero vector in V . Then the $n + 1$ vectors $\alpha, T\alpha, T^2\alpha, \dots, T^n\alpha$ are linearly dependent so that there is a set $\{a_0, a_1, \dots, a_n\}$ of elements of F , not all of them zero, such that

$$a_0\alpha + a_1T\alpha + \dots + a_nT^n\alpha = 0.$$

Writing $g(x) = a_0 + a_1x + \dots + a_nx^n$, we see that $\deg g(x) \leq n$ and $g(T)\alpha = 0$. If $g(T)V = 0$, then, because $m(x)$ is the minimal polynomial, we have $\deg m(x)$

$\leq \deg g(x) \leq n$ and so the theorem holds. Suppose now that $g(T)V \neq 0$. Let $U = \{\alpha: g(T)\alpha = 0\}$. Then U is a proper subspace of V and $\dim U = r < n$. By the inductual assumption there is a polynomial $m_1(x)$ of degree less than or equal to r such that $m_1(T)U = 0$. Moreover, $TU \subseteq U$, since $g(T)$ commutes with T , so that U is an invariant subspace of V . Now consider the quotient space V/U . We have $\dim V/U = n - r < n$ and so, by the inductual assumption again, there is a polynomial $m_2(x)$ of degree $\leq n - r$ such that $m_2(T)V/U = 0$. This means that $m_2(T)V \subseteq U$, so that $m_1(T)m_2(T)V = 0$.

Writing $h(x) = m_1(x)m_2(x)$ we have $h(T)V = 0$ so that if $m(x)$ is the minimal polynomial

$$\deg m(x) \leq \deg h(x) = \deg m_1(x) + \deg m_2(x) \leq r + n - r = n.$$

Thus the theorem holds in all cases and the proof is complete.

NOTE: If $m_1(x)$ and $m_2(x)$ are the minimal polynomials of T restricted to U and V/U respectively, then $h(x) = m_1(x)m_2(x)$ coincides with the minimal polynomial of T .

THEOREM 2. *Let V be a vector space of dimension n over a field F , and let T be a linear transformation on V . If the minimal polynomial $m(x) = (p(x))^s$, where $p(x)$ is irreducible and of degree r , then r divides n .*

Proof. Let $B = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ be a basis of V . We assert that there is one of these vectors, which, with no loss of generality, may be taken to be α_1 , such that the set $G = \{\alpha_1, T\alpha_1, \dots, T^{rs-1}\alpha_1\}$ is linearly independent.

Suppose the statement is false; then for every α in V , $p(T)^j\alpha = 0$ for some $j_\alpha < s$. This is so because the annihilating polynomial of minimal degree (given here by the assumed dependence) of a vector divides any other annihilating polynomial, and in particular, then, divides $(p(x))^s$. Thus $p(T)^{s-1}V = 0$, a contradiction.

Let U be the subspace of V generated by the set G . Then $\dim U$ equals rs . If $U = V$, then $rs = n$ so that r divides n and we are finished. To complete the proof we use an induction on n . Suppose that $U \neq V$. First note that $TU \subseteq U$, since for every $j < rs - 1$, $T(T^j\alpha_1) = T^{j+1}\alpha_1$ is a basis element of U , whereas $(p(T))^s\alpha_1 = 0$ gives $T^{rs}\alpha_1$ in terms of the basis elements. We go now to the quotient space V/U . Since $TU \subseteq U$ we can consider the transformation T_1 induced on V/U by T . Since $(p(T_1))^s$ also annihilates V/U , the minimal polynomial of T_1 is $(p(x))^t$ for some $t \leq s$. Now $\dim V/U = n - rs < n$ so that the inductual hypothesis makes r divide $n - rs$ and this implies that r divides n , completing the proof.

COROLLARY. *Let V be a vector space of dimension n over a field F and let T be a linear transformation on V . If the minimal polynomial $m(x) = (p(x))^s$, where $p(x)$ is irreducible and of degree r , then $\det(xI - T) = (p(x))^{n/r}$.*

Proof. Assume that the statement is true for all spaces of dimension less than n .

The case $n = 1$ is trivial. Let T_1 and T_2 be the linear transformations induced on U and V/U respectively by T . Since $\dim U < n$ and $\dim V/U < n$, the inductive hypothesis gives

$$\det(xI - T_1) = (p(x))^{rs/r} \text{ and } \det(xI - T_2) = (p(x))^{(n-rs)/r}$$

But then

$$\det(xI - T) = \det(xI - T_1)\det(xI - T_2) = (p(x))^{n/r}.$$

REMARK. $\det(xI - T) = f(x)$ is, of course, the characteristic polynomial of T . The corollary extends at once to the following:

THEOREM 3. *Let the minimal polynomial of T be*

$$m(x) = (p_1(x))^{s_1}(p_2(x))^{s_2} \cdots (p_k(x))^{s_k},$$

where each $p_i(x)$ is irreducible of degree r_i . Let V_i be the null space of $(p_i(x))^{s_i}$ and let n_i be the dimension of V_i , then for $i = 1, 2, \dots, k$ we have that r_i divides n_i and

$$f(x) = (p_1(x))^{n_1/r_1} \cdots (p_k(x))^{n_k/r_k}.$$

This is immediate from Theorem 2 and its corollary, since V is the direct sum of the V_i and each $(p_i(x))^{s_i}$ is the minimal polynomial of the transformation T_i induced on V_i by T . By Theorem 1 we have $r_i s_i \leq n_i$ so that $s_i \leq n_i/r_i$ and hence $m(T)V = 0$ implies that $f(T)V = 0$. Thus the Cayley-Hamilton theorem follows as a consequence.

Note that the vector α_1 of Theorem 2 is annihilated by no polynomial of degree $< rs$, and in fact that its order (i.e., annihilating polynomial of minimal degree) is the minimal polynomial $m(x)$. For the general case, in each V_i there is a vector α_i whose order is the minimal polynomial $(p_i(x))^{s_i}$ and the vectors α^j given by $\alpha^j = \alpha_1 + \cdots + \alpha_j$, for $j = 1, 2, \dots, k$, have orders $\prod_{i=1}^j (p_i(x))^{s_i}$. Thus $\alpha = \alpha^k$ has order $m(x)$, the minimal polynomial of T .

ANOTHER PROOF OF THE RATIONAL DECOMPOSITION THEOREM

H. G. JACOB, University of Massachusetts

1. Introduction. The Rational Decomposition Theorem states that a finite dimensional vector space under a linear transformation decomposes into a direct sum of cyclic subspaces. There are at least two rather well-known proofs of this theorem. The more elegant one applies, to the case of a single linear transformation, the theorem that a finitely generated module over a principal ideal domain is the direct sum of cyclic submodules [3, p. 386]. The more direct proof involves showing that a cyclic subspace of maximum dimension is a direct summand in a decomposition

into invariant subspaces [4, p. 309]. The purpose of this note is to give a somewhat different argument for the latter result. It is based on some simple facts about the dual space and the existence of a vector whose minimal polynomial equals that of the linear transformation. The technique used extends that employed by C. W. Curtis in showing that a nilpotent linear transformation yields a decomposition into cyclic subspaces [1, p. 192].

2. Preliminaries. (i) Let T be a linear transformation on a finite dimensional vector space V over a field F . The minimal polynomial $m(x) \in F[x]$ of T needs no explanation [4, p. 306]. The **minimal polynomial** $m_y(x)$ of a vector $y \in V$ relative to T is the monic polynomial of least degree such that $m_y(T)y = 0$. It follows that $m_y(x)$ divides $m(x)$. The existence of $y \in V$ such that $m_y(x) = m(x)$ can be argued in two steps. First, it is immediate when $m(x) = f(x)^e$ where $f(x)$ is an irreducible polynomial and e a nonnegative integer. Second, if $m(x) = f_1(x)^{e_1} f_2(x)^{e_2} \cdots f_r(x)^{e_r}$ where $f_i(x)$ is irreducible then $V = V_1 \oplus V_2 \oplus \cdots \oplus V_r$ where V_i , for $1 \leq i \leq r$, is T -invariant and the minimal polynomial of T restricted to V_i is $f_i(x)^{e_i}$. The vector $y = y_1 + y_2 + \cdots + y_r$ where $y_i \in V_i$ and $m_{y_i}(x) = f_i(x)^{e_i}$ is the desired element. The theorem that V decomposes into the direct sum of T -invariant space V_i is sometimes called the Primary Decomposition Theorem [2, p. 220].

(ii) The space V^* of all linear forms on V , linear transformations from V to F , is called the **dual space of v** . For $y^* \in V^*$ and $y \in V$ we use the symbol $\langle y^*, y \rangle$ to denote $y^*(y) \in F$. Thus

$$\langle y_1^* + y_2^*, y \rangle = \langle y_1^*, y \rangle + \langle y_2^*, y \rangle \text{ for } y_1^* \text{ and } y_2^* \text{ in } V^* \text{ and } y \text{ in } V.$$

$$\langle y^*, y_1 + y_2 \rangle = \langle y^*, y_1 \rangle + \langle y^*, y_2 \rangle \text{ for } y^* \in V^* \text{ and } y_1 \text{ and } y_2 \text{ in } V.$$

$$\langle \alpha y^*, y \rangle = \alpha \langle y^*, y \rangle = \langle y^*, \alpha y \rangle \text{ for } y^* \in V^* \text{ and } y \in V \text{ and } \alpha \in F.$$

By associating y to the linear form $\langle y^*, y \rangle$ on V^* we can identify V naturally with $V^{**} = (V^*)^*$. Moreover, for a basis y_1, y_2, \dots, y_n of V there exists a basis $y_1^*, y_2^*, \dots, y_n^*$, called the dual basis of y_1, y_2, \dots, y_n such that

$$\langle y_i^*, y_j \rangle = \delta_{ij}, \text{ the Kronecker delta.}$$

Since $\langle y^*, Ty \rangle$, for y^* fixed, is a linear form there exists an element denoted by T^*y^* in V^* such that

$$\langle T^*y^*, y \rangle = \langle y^*, Ty \rangle.$$

It is readily seen that the mapping T^* which maps y^* to T^*y^* is a linear transformation on V^* . If $f(x) \in F[x]$ then

$$\langle f(T^*)y^*, y \rangle = \langle y^*, f(T)y \rangle.$$

From this it follows that T and T^* have the same minimal polynomial.

(iii) Let W be a subspace of V with $\dim W = r$ and define

$$W^\perp = \{y^* \in V^* \mid \langle y^*, y \rangle = 0 \text{ for all } y \in W\}.$$

Then W^\perp is a subspace of V^* with $\dim W^\perp = \dim V - r$ and $W^{\perp\perp} = (W^\perp)^\perp = W$. Furthermore, if W is T -invariant then W^\perp is T^* -invariant. Consequently

$$\begin{aligned} V &= W_1 \oplus W_2 \text{ with } W_i \text{ } T\text{-invariant} \\ \Rightarrow V^* &= W_1^\perp \oplus W_2^\perp \text{ with } W_i^\perp \text{ } T^*\text{-invariant.} \end{aligned}$$

Of course the dual also holds, i.e.,

$$\begin{aligned} V^* &= W_1^* \oplus W_2^* \text{ with } W_i^* \text{ } T^*\text{-invariant} \\ \Rightarrow V &= W_1^{*\perp} \oplus W_2^{*\perp} \text{ with } W_i^{*\perp} \text{ } T\text{-invariant.} \end{aligned}$$

(iv) A subspace W of V is said to be **T-cyclic** if there exists a vector $y \in W$ and a nonnegative integer r such that $y, Ty, \dots, T^r y$ form a basis for W . Thus for the vector y if the degree of $m_y(x)$ is k then $y, Ty, \dots, T^{k-1}y$ are linearly independent and the space W spanned by these k vectors is T -cyclic.

3. Main Theorem. We are now in a position to prove our principal result.

THEOREM: Let V be an n -dimensional vector space (n finite) and T a linear transformation on V . Then V is the direct sum of T -cyclic subspaces.

Proof. Let k be the degree of the minimal polynomial $m(x)$ of T , and let y be a vector in V with $m_y(x) = m(x)$. Then the space W spanned by $y, Ty, \dots, T^{k-1}y$ is T -cyclic. We shall prove that if $W \neq V$ ($k \neq n$) then there exists a T -invariant subspace W' such that $V = W \oplus W'$. Clearly, by induction on the dimension, W' will then be the direct sum of T -cyclic subspaces and the proof complete.

To show the existence of W' enlarge the basis $y_1 = y, y_2 = Ty, \dots, y_k = T^{k-1}y$ of W to a basis $y_1, y_2, \dots, y_k, \dots, y_n$ of V and let $y_1^*, y_2^*, \dots, y_k^*, \dots, y_n^*$ be the dual basis. To simplify notation let $y^* = y_k^*$. Then

$$\langle y^*, y_i \rangle = 0 \text{ for } 1 \leq i \leq k-1 \text{ and } \langle y^*, y_k \rangle = 1.$$

Consider the space W^* spanned by $y^*, T^*y^*, \dots, T^{*k-1}y^*$. Since $m(x)$ is also the minimal polynomial of T^* the space W^* is T^* -invariant. Now observe that if $W^* \cap W^\perp = \{0\}$ and $\dim W^* = k$ then $V^* = W^* \oplus W^\perp$ where W^* and W^\perp are T^* -invariant (since $\dim W^\perp = n - k$). This in turn implies (from iii) the desired decomposition $V = W^{\perp\perp} \oplus W^{*\perp} = W \oplus W'$ where $W^{\perp\perp} = W$ and $W^{*\perp} = W'$ are T -invariant.

Finally we shall prove that $W^* \cap W^\perp = \{0\}$ and $\dim W^* = k$ simultaneously as follows. Suppose that $a_0 y^* + a_1 T^* y^* + \dots + a_s T^{*s} y^* \in W^\perp$ where $a_s \neq 0$ and $0 \leq s \leq k-1$. Then

$$\begin{aligned} T^{*k-1-s} (a_0 y^* + a_1 T^* y^* + \dots + a_s T^{*s} y^*) \\ = a_0 T^{*k-1-s} y^* + a_1 T^{*k-s} y^* + \dots + a_s T^{*k-1} y^* \end{aligned}$$

is in W^\perp since W^\perp is T^* -invariant. Therefore

$$\langle (a_0 T^{*k-1-s} + a_1 T^{*k-s} + \cdots + a_s T^{*k-1}) y^*, y \rangle = 0.$$

This implies (from ii)

$$\begin{aligned} & \langle y^*, (a_0 T^{k-1-s} + a_1 T^{k-s} + \cdots + a_s T^{k-1}) y \rangle \\ &= a_0 \langle y^*, y_{k-s} \rangle + a_1 \langle y^*, y_{k-s+1} \rangle + \cdots + a_s \langle y^*, y_k \rangle = a_s = 0 \end{aligned}$$

which is a contradiction.

References

1. C. W. Curtis, *Linear Algebra*, Allyn and Bacon, Boston, 1968.
2. K. Hoffman and R. Kunze, *Linear Algebra*, Prentice Hall, Englewood Cliffs, N. J., 1971.
3. S. Lang, *Algebra*, Addison Wesley, Reading, Mass., 1965.
4. L. J. Page and J. D. Swift, *Elements of Linear Algebra*, Blaisdell, Waltham, Mass., 1961.

A PROOF OF THE CHAIN RULE FOR DERIVATIVES IN n -SPACE

A. G. FADELL, State University of New York at Buffalo

The usual proofs of the chain rule for derivatives in n -space utilize an $\varepsilon - \delta$ argument by approximations. In the view of the author the following proof is of the kind the average student finds easiest to follow. We use the notation of R. G. Bartle's *The Elements of Real Analysis*, Wiley, 1964, Sec. 20.

We are given a function f with domain $D(f)$ in R^q and values in R^r , and g a function with domain $D(g)$ in R^p and range in R^q , with $g(c) = a$, where a is an interior point of $D(f)$ and c an interior point of $D(g)$.

Assume that f has a derivative L_f at $u = a$ and that g has a derivative L_g at $x = c$. We show that the composition $f \circ g$ has a derivative $L_f \circ L_g$ at $x = c$. Let

$$(1) \quad \phi_f(u) = \begin{cases} \frac{f(u) - f(a) - L_f(u - a)}{|u - a|}, & u \neq a \\ 0, & u = a. \end{cases}$$

Then ϕ_f is continuous at $u = a$, and in view of (1) for any $u \in D(f)$

$$(2)' \quad f(u) - f(a) = \phi_f(u) |u - a| + L_f(u - a).$$

Letting $u = g(x)$, $a = g(c)$ in (2) and subtracting $L_f[L_g(x-c)]$ from both sides we have

$$f[g(x)] - f[g(c)] - L_f[L_g(x-c)] = \phi_f[g(x)] |g(x) - g(c)| + L_f[g(x) - g(c)] - L_f[L_g(x-c)].$$

Dividing both sides by $|x - c|$ and using the linearity of L_f we obtain

$$\frac{f[g(x)] - f[g(c)] - L_f[L_g(x-c)]}{|x-c|} = \phi_f[g(x)] \frac{|g(x) - g(c)|}{|x-c|} + L_f \left[\frac{g(x) - g(c) - L_g(x-c)}{|x-c|} \right].$$

Finally, the right side has limit 0 as $x \rightarrow c$ since $\phi_f[g(x)] \rightarrow \phi_f[g(c)] = \phi_f(a) = 0$, $|g(x) - g(c)|/|x - c|$ is locally bounded, and the last term has limit $L_f(0) = 0$, since g is differentiable at c . Thus, by definition $L_f \circ L_g$ is the derivative of $f \circ g$.

MATHEMATICAL EDUCATION

EDITED BY J. G. HARVEY AND M. W. POWNALL

Material for this Department should be sent to Shirley Hill, Department of Mathematics, University of Missouri, Kansas City, MO 64110, or to Paul Mielke, Department of Mathematics, Wabash College, Crawfordsville, IN 47933.

SURVIVAL OF THE TWO-YEAR COLLEGE MATHEMATICS TEACHER

P. A. LINDSTROM, Genesee Community College, Batavia, New York

In "Survival kit for the college mathematician" [4] and "Survival for mathematicians or mathematics?" [6] both Professors Flanders and Peterson discuss the future of teaching and research at the college and university levels. Flanders believes that college teachers of mathematics have a professional obligation to survive as mathematicians. Peterson believes that the problem of survival affects not only the level of college mathematicians, but all levels of the academic mathematical system.

How then is the teacher of mathematics at the two-year college to survive? Many people would say that since the two-year college faculties are highly student-oriented, more so than the discipline-oriented faculties of the four-year colleges and universities, then the two-year college teacher of mathematics can survive by being a "good teacher." It is not the purpose of this paper to dwell upon his survival by being a "good teacher." Instead, there are two other areas of survival that affect him and that are closely related to that of being a "good teacher." These are the overlapping areas of 1. professional obligations and 2. professional identity.

Teaching his classes, advising students, serving on committees, etc. are but some of the obligations of the two-year college mathematics teacher. With regard to his profession he also has many obligations so that he can survive as a mathematician and a teacher. Christie and Wells [1], Flanders [4], and various CUPM publications [2 and 3], suggest many ways that are applicable to the two-year college mathematics teacher. Some of these are:

1. Scan and read textbooks and journals to keep up to date with mathematics and to find new and interesting problems and material for the classroom.
2. Organize a mathematics discussion group to discuss not only possible new courses, teaching techniques, textbook selection, etc., but also journal articles, mathe-

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

ASSOCIATE EDITORS: JOSHUA BARLAZ, ERIC S. LANGFORD. COLLABORATING EDITORS: LEONARD CARLITZ, GULBANK D. CHAKERIAN, HASKELL COHEN, S. ASHBY FOOTE, ISRAEL N. HERSTEIN, MURRAY S. KLAMKIN, DANIEL J. KLEITMAN, ROGER C. LYNDON, MARVIN MARCUS, CHRISTOPH NEUGEBAUER, ALBERT WILANSKY, AND UNIVERSITY OF MAINE PROBLEMS GROUP: EARL M. L. BEARD, GEORGE S. CUNNINGHAM, CLAYTON W. DODGE, OSKAR FEICHTINGER, WILLIAM R. GEIGER, RAMESH GUPTA, GARY HAGGARD, PHILIP M. LOCKE, JOHN C. MAIRHUBER, CURTIS S. MORSE, GRATTAN P. MURPHY, EDWARD S. NORTHAM AND WILLIAM L. SOULE, JR.

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave. Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Mathematics Department, University of Maine, Orono, ME 04473. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before March 31, 1974.

E 2444. *Proposed by Ray Redheffer, University of California, Los Angeles*

Let S be an open connected subset of real Euclidean space R^n and suppose that $f: S \rightarrow R^m$ is differentiable. Let the Jacobian matrix $Df(x)$ at $x \in S$ satisfy

$$\|Df(x)\| \leq \sigma \|f(x)\|$$

for some constant σ and all $x \in S$, where the norm of a matrix is the sum of the absolute values of its entries. Show that if $x_1, x_2 \in S$ can be connected by a path of length d lying wholly within S , then

$$\|f(x_1)\| \leq \|f(x_2)\| e^{\sigma d}.$$

(Note that a consequence of this is that f cannot vanish anywhere on S unless it vanishes everywhere on S .)

E2445. *Proposed by F. Leuenberger, Feldmeilen, Switzerland*

Let P be a point in the interior of a triangle ABC . Let R_1, R_2, R_3 denote the distances from P to the vertices of ABC and let r_1, r_2, r_3 denote the perpendicular distances from P to the sides of ABC . Show that

$$\sum \frac{r_2 + r_3}{r_2 + 2R_1 + r_3} \leq 1 \leq \frac{1}{3} \sum \frac{R_1}{r_2 + r_3}$$

with equality if and only if the triangle is equilateral and P is its center.

E 2446. *Proposed by H. D. Ruderman, Hunter College High School*

Characterize those moduli m for which both $x^3 \equiv 1 \pmod{m}$ implies $x \equiv 1 \pmod{m}$ and $x^3 \equiv 0 \pmod{m}$ implies $x \equiv 0 \pmod{m}$; show that these are precisely the moduli for which $x^3 \equiv y^3 \pmod{m}$ implies $x \equiv y \pmod{m}$ for all x, y .

E 2447. *Proposed by E. T. H. Wang, University of Waterloo, Canada*

A k -satisfactory sequence is a k -tuple $S = (a_1, a_2, \dots, a_k)$ of natural numbers with $a_1 \leq a_2 \leq \dots \leq a_k$ such that $\sum a_i = \prod a_i$. Let $v(S)$ denote this common value. Show that $v(S) \leq 2k$ with equality if and only if $S = (1, \dots, 1, 2, k)$, and investigate the problem of finding a lower bound for $v(S)$. (Cf. E 2262 [1971, 1021].)

E 2448. *Proposed by Gérard Letac, Université de Clermont, France*

Find all positive semi-definite Hermitian matrices $A = (a_{ij})$ with the property that the matrix of reciprocals $(1/a_{ij})$ is also positive semi-definite.

E 2449. *Proposed by Frank Siwiec, John Jay College*

Let f be a continuous mapping of \mathbb{R} onto \mathbb{R} with the property that for every $y \in \mathbb{R}$, the boundary of the set $f^{-1}(y) = \{x \in \mathbb{R} : f(x) = y\}$ is compact. Show that f is a closed mapping.

SOLUTIONS OF ELEMENTARY PROBLEMS

A Curious Summation Inequality

E 2373 [1972, 905]. *Proposed by Grahame Bennett, Indiana University*

Let r_1, r_2, \dots, r_n be real numbers. Show that there exists a subset N of $\{1, 2, \dots, n\}$, neither containing nor omitting three consecutive integers, such that

$$\left| \sum_{j \in N} r_j \right| \geq \frac{1}{6} \sum_{j=1}^n |r_j|.$$

Show further that $1/6$ is the best possible constant here.

Establish the corresponding result (with $1/6$ replaced by $1/3\pi$) for complex numbers.

Solution by L. E. Mattics, University of South Alabama and C. S. Gardner, University of Texas (independently). Let $\sum |r_j| = s$ and for $i = 0, 1, 2$, let $p(i) = \sum \{r_j : r_j \geq 0 \text{ and } j \equiv i \pmod{3}\}$ and $n(i) = \sum \{r_j : r_j < 0 \text{ and } j \equiv i \pmod{3}\}$. There exist distinct i, i' , such that either $p(i) + p(i') \geq s/3$ or such that $n(i) + n(i') \leq -s/3$; we assume without loss of generality that the first case holds. Now if $p(i) + p(i') \geq -n(i) - n(i')$ then $2(p(i) + p(i')) + n(i) + n(i') \geq s/3$, so that either $p(i) + p(i') + n(i) \geq s/6$ or $p(i) + p(i') + n(i') \geq s/6$. Similarly if

$p(i) + p(i') \leq -n(i) - n(i')$ it follows that either $-p(i) - n(i) - n(i') \geq s/6$ or $-p(i') - n(i) - n(i') \geq s/6$. This establishes the inequality. To show that the constant $1/6$ is best possible, take $r_1 = r_2 = r_3 = 1$ and $r_4 = r_5 = r_6 = -1$.

Now suppose that complex numbers z_1, \dots, z_n are given. Write $z_j = r_j \exp(i\theta_j)$ and set $\sum |z_j| = \sum r_j = s$. Let $F(\theta) = \sum r_j |\cos(\theta_j - \theta)|$; an elementary computation shows that

$$\int_0^{2\pi} F(\theta) d\theta = 4s,$$

so by the Mean Value Theorem, there exists θ_0 such that $F(\theta_0) = 2s/\pi$. Applying the inequality derived above for real numbers, we see that there exists a subset $N \subseteq \{1, 2, \dots, n\}$ of the desired type such that

$$\left| \sum_{j \in N} r_j \cos(\theta_j - \theta_0) \right| \geq \frac{1}{6} F(\theta_0) = \frac{s}{3\pi}.$$

From this, it follows that

$$\begin{aligned} \frac{s}{3\pi} &\leq \left| \sum_{j \in N} r_j \cos(\theta_j - \theta_0) \right| = \left| \sum_{j \in N} \operatorname{Re}(r_j e^{i(\theta_j - \theta_0)}) \right| \\ &= \left| \operatorname{Re}(e^{-i\theta_0} \sum_{j \in N} z_j) \right| \leq \left| e^{-i\theta_0} \sum_{j \in N} z_j \right| = \left| \sum_{j \in N} z_j \right| \end{aligned}$$

establishing the proof of the inequality in the complex case. To show that $1/3\pi$ is actually best possible, let $n = 12m$ and consider the sequence $\omega_1, \omega_2, \dots, \omega_n$ of the n th roots of unity, where $\omega_j = \exp(i\theta_j)$ and $\theta_j = 2\pi j/n$. Let $N \subseteq \{1, 2, \dots, n\}$ be any subset of the specified type. Write $\omega = \sum_{j \in N} \omega_j$ and let $\theta = \arg \omega$. Note that

$$\begin{aligned} \left| \sum_{j \in N} \cos(\theta_j - \theta) \right| &= \left| \sum_{j \in N} \operatorname{Re}(e^{i(\theta_j - \theta)}) \right| \\ &= \left| \operatorname{Re}(e^{-i\theta} \sum_{j \in N} \omega_j) \right| = \left| \operatorname{Re}(e^{-i\theta} |\omega| e^{i\theta}) \right| \\ &= |\omega| = \left| \sum_{j \in N} \omega_j \right|. \end{aligned}$$

Now partition the unit circle into $4m$ half-open subintervals I_1, I_2, \dots, I_{4m} , where

$$I_k = \left\{ z : |z| = 1 \text{ and } (k-1)\frac{\pi}{2m} < \arg z \leq \frac{k\pi}{2m} \right\}.$$

Note that by assumption on N , the number of ω_j with $j \in N$ which lie in any I_k is either one or two—never zero or three (or more). Moreover, if we translate the $\omega_j \bmod 2\pi$ by considering $\omega'_j = e^{-i\theta} \omega_j$, then the same will hold true for the ω'_j

except for a possible "edge effect" which is negligible for large n . (More precisely, one (and only one) I_k could contain zero or three ω'_j because of the fact that possibly N contains (or excludes) n , 1, and 2 or $n-1$, n , and 1.) Now

$$\left| \sum_{j \in N} \omega_j \right| = \left| \sum_{j \in N} \cos(\theta_j - \theta) \right| = \left| \sum_1 \cos(\theta_j - \theta) + \sum_2 \cos(\theta_j - \theta) \right|,$$

where the first sum is over all $j \in N$ with $\cos(\theta_j - \theta) > 0$ and the second over all $j \in N$ with $\cos(\theta_j - \theta) < 0$; that is, the first sum is over all $j \in N$ such that $\omega'_j \in I_1 \cup \dots \cup I_m \cup I_{3m+1} \cup \dots \cup I_{4m}$ and the second over all $j \in N$ such that $\omega'_j \in I_{m+1} \cup \dots \cup I_{3m}$. (Those θ_j such that $\cos(\theta_j - \theta) = 0$ can be ignored.) Assume without loss of generality that $\sum_1 + \sum_2 \geq 0$. Using the fact that for large n , the sums can be approximated by integrals, and the facts that in every subinterval in \sum_1 there are at most two ω'_j and that in every subinterval in \sum_2 there is at least one ω'_j we see that

$$\begin{aligned} \left| \sum_{j \in N} \omega_j \right| &= \sum_1 \cos(\theta_j - \theta) + \sum_2 \cos(\theta_j - \theta) \\ &\leq \left(\frac{2m}{\pi} \right) \int_{-\pi/2}^{\pi/2} 2 \cos \theta' d\theta' + \left(\frac{2m}{\pi} \right) \int_{\pi/2}^{3\pi/2} \cos \theta' d\theta' + o(1) \\ &= \frac{4m}{\pi} + o(1) = \frac{n}{3\pi} + o(1) = \frac{1}{3\pi} \sum_{j=1}^n |\omega_j| + o(1). \end{aligned}$$

This shows that $1/3\pi$ is best possible in the complex case.

Also solved by the proposer.

Editor's comment. Related inequalities appear in the literature, but without the added feature of "neither containing nor omitting three consecutive integers." The proposer refers to the following inequality

$$(*) \quad \left| \sum_{j \in X} z_j \right| \geq \frac{1}{\pi} \sum_{j=1}^n |z_j|$$

for complex numbers which can be found in Bourbaki, *General Topology* (part 2), Addison-Wesley, 1966, Chap. VIII, Ex, 1, § 3, p. 126. Bourbaki notes that the constant $1/\pi$ is best possible, but that it cannot be achieved. This leads one to believe that the constant $1/3\pi$ for our problem, although best possible, cannot be achieved as can the constant $1/6$ for the real case. Edwin Klein calls attention to the inequality (*) with the constant $1/6$ (rather than $1/\pi$) which is derived in Rudin, *Real and Complex Analysis*, McGraw-Hill, New York, 1966, p. 119. Rudin's argument is simpler than Bourbaki's which is to be expected since his constant is less precise.

A Birthday Problem

E 2386 [1972, 1134]. Proposed by William Knight, University of New Brunswick

The classical birthday problem can be phrased as a bet between a statistics teacher

and a class of $n < 365$ students, the teacher betting that at least two students have the same birthday. (The usual stake is one-up-ness rather than money.) If birthdays are (1) independently and (2) uniformly distributed over the 365 days of the year (leap years being ignored) the probability of the teacher's winning is $1 - (365)_n/365^n$ where $(m)_n$ denotes the partial factorial $m!/(m-n)!$. But it is more likely that birthdays are not really equally numerous at all seasons. Show that this, in fact, makes the bet more favorable for the teacher; that is, if assumption (2) is dropped, $1 - (365)_n/365^n$ is a lower bound attained only when all days are equally probable as birthdays.

Solution by D. M. Bloom, Brooklyn College. The assumption $n \geq 2$ is clearly intended; also, we may assume by induction that the result is true for all "years" of fewer than 365 days (the result being trivial for a year of just one day). Let x_i be the probability of a birthday occurring on the i th day of the year; then the probability that the teacher loses is

$$P(x_1, \dots, x_{365}) = (n!) \sum_S \left(\prod_{i \in S} x_i \right)$$

where S runs over all n -element subsets of $\{1, \dots, 365\}$. Since P is continuous, the maximum of P on the closed set $\{x_i \geq 0 \text{ (all } i), \sum x_i = 1\}$ exists, and it cannot occur when any x_i is zero (by the induction hypothesis and the fact that $(364)_n/364^n < (365)_n/365^n$); hence the method of Lagrange multipliers is applicable (with $\sum x_i = 1$ as the side condition) and implies that $\partial P/\partial x_i = \partial P/\partial x_j$ for all i, j at the point in question. Therefore

$$(*) \quad 0 = \partial P/\partial x_i - \partial P/\partial x_j = (n!)(x_j - x_i) \sum_T \left(\prod_{k \in T} x_k \right),$$

where T runs over all $(n-2)$ -element subsets of $\{1, \dots, 365\}$ which contain neither i nor j . Since $n \geq 2$, the summation in $(*)$ is nonempty and hence nonzero, so that $x_j - x_i = 0$, $x_i = x_j$ (all i, j) which is the desired result.

Also solved by Ellen Hertz, Harry Lass, Carolyn MacDonald, William Nuesslein, G. S. Rogers & D. L. Young, Michael Shimshoni (Israel), and the proposer.

Editorial Note. Various versions of the birthday problem have been dealt with in the literature. The reader is referred to articles on this subject in the *American Statistician*, Feb. 1968, April 1968, Feb. 1970, June 1972.

A Pair of Triangle Inequalities

E 2388 [1972, 1135]. *Proposed by A. W. Walker, Toronto, Canada*

Let a, b, c ; s, r, R, I, H denote the side lengths, semiperimeter, inradius, circumradius, incenter and orthocenter of a triangle ABC .

(i) For ABC arbitrary, prove that

$$bc + ca + ab \geq (AI + BI + CI)^2$$

with equality if and only if the triangle is equilateral.

(ii) For ABC non-obtuse, prove that $s^2 \geq 2R^2 + 8Rr + 3r^2$ or, equivalently,

$$a^2 + b^2 + c^2 \geq (AH + BH + CH)^2,$$

with equality if and only if ABC is equilateral or right isosceles.

Solution by M. G. Greening, University of New South Wales, Australia.

$$\begin{aligned} \text{(i)} \quad \Sigma ab - (\Sigma AI)^2 &= 4R^2 [\Sigma \sin A \sin B \\ &\quad - 4(\Sigma \sin^2 \tfrac{1}{2}A \sin^2 \tfrac{1}{2}B + 2 \sin \tfrac{1}{2}A \sin \tfrac{1}{2}B \sin \tfrac{1}{2}C \Sigma \sin \tfrac{1}{2}A)] \\ &= 16R^2 [3 \sin \tfrac{1}{2}A \sin \tfrac{1}{2}B \sin \tfrac{1}{2}C - 2 \sin \tfrac{1}{2}A \sin \tfrac{1}{2}B \sin \tfrac{1}{2}C \Sigma \sin \tfrac{1}{2}A] \\ &= 4Rr(3 - 2 \Sigma \sin \tfrac{1}{2}A) \geq 0, \quad (\text{See [1], 2.10.}) \end{aligned}$$

with equality holding only for ABC equilateral.

(ii) Set $\alpha = \pi - 2A$, $\beta = \pi - 2B$, $\gamma = \pi - 2C$. Then

$$\Sigma a^2 - (\Sigma AH)^2 = 4R^2(\Sigma \cos^2 \tfrac{1}{2}\alpha - [\Sigma \sin \tfrac{1}{2}\alpha]^2) \geq 0$$

and this is (ii). See [2].

If $A = \tfrac{1}{2}\pi$, then $\Sigma \cos^2 \tfrac{1}{2}\alpha = 2$ and $(\Sigma \sin \tfrac{1}{2}\alpha)^2 = 1 + \sin \beta$, so that equality holds only for $\tfrac{1}{2}\beta = \tfrac{1}{4}\pi = \tfrac{1}{2}\gamma$.

Otherwise (ii) depends on the inequalities

(iii) $(\sin \tfrac{1}{2}\alpha - \tfrac{1}{2})(\sin \tfrac{1}{2}\beta - \tfrac{1}{2}) \geq 0$ and

(iv) $2 \sin \tfrac{1}{2}\beta \sin \tfrac{1}{2}\gamma + \sin \tfrac{1}{2}\alpha \leq 1$

which is equivalent to $\cos \tfrac{1}{2}(\beta - \gamma) \leq 1$, so that equality holds only for ABC equilateral.

The resulting inequality $s^2 \geq 2R^2 + 8Rr + 3r^2$ is stronger than either $s^2 \geq 3r(4R + r)$ or $s^2 \geq (16R - 5r)r$ mentioned in [1].

[1] Bottema et al., *Geometric Inequalities*.

[2] Problem E 1272, this MONTHLY, 67 (1960) 693-694.

Also solved by Anders Bager (Denmark), Leon Bankoff, A. G. Ferrer (Mexico), C. S. Gardner, Leon Gerber, Leonard Goldstone, M. S. Klamkin, and the proposer.

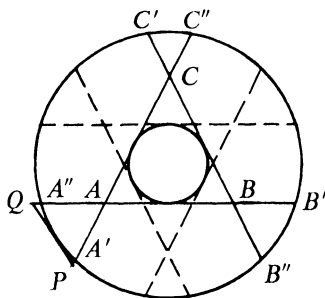
Six Equal Regions? Yes. Seven? No.

E'2391 [1973, 74]. *Proposed by V. R. R. Uppuluri, Oak Ridge National Laboratory*

It is well known that three chords can divide a circular disk into at most seven pieces. Can these seven pieces all have the same area?

I. *Solution by the Rose-Hulman Problems Group.* The answer is no. Each of the

three chords must intersect the other two and divide the disk into two sections with areas of ratio four to three. The envelope of all such chords is a circle. We construct the figure below.



Let $A'B'$ divide the disk into a ratio of four to three. If $C'B'$ is also such a chord, one easily verifies that the angle $A''BC'$ is acute. By symmetry, the construction of the final chord, $A'C''$, yields an equilateral triangle, ABC . After rotating through π radians we find that the image of the region $CC''C'$ (broken lines) is properly contained in the region $A'B''BA$ and therefore $CC''C'$ and $A'B''BA$ cannot have equal areas.

II. *Solution by D. W. Atkinson, University of Nebraska at Omaha.* The answer is no. Assume a solution does exist. It is easy to show that this solution must be symmetrical; i.e., all chords intersect in 60° angles and are equidistant from the center of the circle. Let s be the distance from the center to each chord. The area of the central equilateral triangle is $3s^2\sqrt{3}$. In a unit circle we have $3s^2\sqrt{3} = \pi/7$ or $s = [\pi/(21\sqrt{3})]^{1/2}$. This should also be the area of one of the "triangle-like" pieces ($AA'A''$ in the figure). Construct a tangent to the circle and extend the two chords forming the "triangle-like" piece to form an equilateral triangle (APQ in the figure) which contains the "triangle-like" piece. It should have area greater than $\pi/7$. However, its height is $1 - 2s$, and so its area is $(1 - 2s)^2/\sqrt{3}$. Substituting for s shows that this area is less than $\pi/7$. Thus no solution exists.

III. *Solution by V. Linis, University of Ottawa.* The answer is no. In the seven piece configuration each of the three chords divides the area of the disk in the ratio 3:4. Such a chord subtends a central angle α which satisfies the equation

$$(*) \quad \alpha - \sin \alpha = 6\pi/7$$

and has distance $d = r \cos \frac{1}{2}\alpha$ from the center (r = radius of the disk). If we let $\alpha = \pi - 2\beta$ the equation $(*)$ becomes $2\beta + \sin 2\beta = \pi/7$ which has an approximate solution $\beta = \pi/28$ since β is rather small. Then $d = r \sin \beta = \pi r/28$. The central piece is a triangle with d as inradius, therefore its semiperimeter s satisfies the equation $sd = \pi r^2/7$. It follows that $s = 4r$, which is plainly impossible.

IV. *Comment by M. S. Klamkin, Ford Motor Company (similar comment by R. C. Buck, University of Wisconsin at Madison).* The answer is negative even if the

circular region is replaced by a convex region; see R. C. Buck and E. F. Buck, *Equipartition of convex sets*, Math. Mag., 22 (1949) 195–198, where it is shown that at most six of the regions can have the same area and that these equal regions must be the six outer ones.

Also solved by R. P. A'Hern (England), Peter Avery (England), Anders Bager (Denmark), Merrill Barnebey, C. C. Clever & K. L. Yocom, R. B. Eggleton, Arthur Gittleman, Michael Goldberg, S. H. Greene, Ralph Jones, L. L. Keener, Dan Kenway & Rici Liknaitzky, P. G. Kirmser, Lew Kowarski, O. P. Lossers (Netherlands), Carolyn MacDonald, Carl Maltz, Greg Maxwell, M. D. Meyerson, Larry Olson, C. C. Oursler, W. W. Parsons, D. B. Price, E. S. Rosenthal, Ralph Seifert, Jr., Phil Tracy, G. Tsintsifas (Greece), J. H. Wahab, K. G. Willett, and J. N. Younglove.

Volume of a Simplex

E2393 [1973, 75]. *Proposed by M. S. Klamkin, Ford Motor Company*

Parallel lines are drawn through the vertices A_0, A_1, \dots, A_n of a given simplex of volume V , terminating in the opposite faces (extended if necessary) in the points B_0, B_1, \dots, B_n , respectively.

(1) Show that the volume of the simplex determined by B_0, B_1, \dots, B_n is nV .

(2) Show that the volume of the simplex determined by the vertices $A_0, A_1, \dots, A_r, B_{r+1}, B_{r+2}, \dots, B_n$ is given by $V'_r = |n - r - 1| V$.

Solution by Leon Gerber, St. John's University. Parallel lines are drawn through the vertices A_0, A_1, \dots, A_n , etc. Let the weights of the point P with respect to the given simplex be (p_i) where $\sum_{i=0}^n p_i = s$, with $s = 1$ if P is a proper point, and $s = 0$ if P is improper. Then the cevians A_iP (which are parallel if P is improper) meet the face opposite A_i in $B_i = (b_{ij})$ where $b_{ii} = 0$ and $b_{ij} = p_j/(s - p_i)$. Thus the ratio of the content of $A_0 \cdots A_{r-1} B_r \cdots B_n$ to that of the given simplex is

$$\begin{aligned} & \det \begin{vmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ b_{r0} & \cdots & b_{r,r-1} & b_{r,r} & \cdots & b_{r,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ b_{n,0} & \cdots & b_{n,r-1} & b_{n,r} & \cdots & b_{n,n} \end{vmatrix} \\ &= \det \begin{vmatrix} 0 & \cdots & p_n/(s - p_r) \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ p_r/(s - p_n) & \cdots & 0 \end{vmatrix} \\ &= (r - n) \prod_{i=r}^n p_i/(p_i - s). \end{aligned}$$

The absolute value is $n - r$ if $s = 0$.

Also solved by G. Tsintsifas (Greece), and the proposer. M. G. Greening (Australia) submitted a very clear solution to the first part.

A Combinatorial Problem

E 2395 [1973, 75]. *Proposed by H. W. Gould, West Virginia University*

Let n be a nonnegative integer, For $p = 1, 2, \dots$ define

$$A_p(n) = \sum_{0 \leq k \leq n/2} (-1)^k \left\{ \binom{n}{k} - \binom{n}{k-1} \right\}^p,$$

where we make the usual conventions regarding binomial coefficients. Prove that, whenever n is odd, $A_2(n) = nA_1(n)$.

Solution by the St. Olaf Problem Group. Let $n = 2m + 1$. Equating the coefficients of x^{2m} in $(1+x)^n(1-x)^n = (1-x^2)^n$ and simplifying, one finds that $A_2(n) = (-1)^m \binom{n}{m}$. Also if we note that $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$, it follows that

$$A_1(n) = (-1)^m \left[\binom{n-1}{m} - \binom{n-1}{m-1} \right] = \frac{(-1)^m}{n} \binom{n}{m}.$$

Hence $A_2(n) = nA_1(n)$.

Also solved by M. T. Bird, D. M. Bloom, Robert Breusch, R. A. Gibbs & H. S. Stocker, Elliot Goldstein & Robert Spira, M. G. Greening (Australia), O. P. Lossers (Netherlands), Joseph O'Rourke, Phil Tracy, and David Zeitlin.

Editorial Comment. The result is not true if n is even, since in that case $A_1(4) = 0$ and $A_2(4) \neq 0$. It would be of interest to find some kind of recurrence relation between $A_p(n)$ and $A_{p-1}(n)$ valid for any p .

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers—The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate, signed sheets and should be mailed before March 31, 1974.

An asterisk () means neither the proposer nor the editors supplied a solution.*

5940. *Proposed by Donald Minassian, Indianapolis, Indiana*

Let R be a commutative ring with 1 and let I be an ideal of R . On p. 131 of *Introduction to Modern Algebra*, (D. C. Heath, 1963) W. Barnes claims I is noetherian (i.e., as a ring which may lack a unit, and defining the ideal generated by a subset T of a ring as the smallest ideal containing T). Prove or give a counterexample.

5941. *Proposed by Jan Mycielski, University of Colorado*

Prove (without using the axiom of choice) that \mathbf{R}/\mathbf{Q} is of the same cardinality as \mathbf{B}/\mathbf{F} , where \mathbf{R} is the additive group of real numbers, \mathbf{Q} is the additive group of rational

numbers, \mathbf{B} is the Boolean algebra of all subsets of the set of integers, and \mathbf{F} is the ideal of finite sets of integers.

5942*. *Proposed by D. M. Bloom, Brooklyn College*

Let X_1, X_2, X_3 be independent random variables such that $E(X_1) > E(X_2) > E(X_3)$. Assume that the X_i have normal distributions with a common variance. Prove or disprove: if $P(X_1 > X_2) = K$ and $P(X_2 > X_3) = L$, then $P(X_1 > X_3)$ is greater than M where M is defined by

$$\frac{K}{1-K} \cdot \frac{L}{1-L} = \frac{M}{1-M}.$$

5943. *Proposed by L. J. Wallen, University of Hawaii*

Let V be a vector space over some field and let $L(V)$ denote the algebra of all endomorphisms of V . A set $\phi \subset L(V)$ is transitive if $\phi x = V$ for each $x \in V$, $x \neq 0$. Let Ω be a transitive subalgebra of $L(V)$. Determine the automorphisms α of Ω having the property that whenever $\phi \subset \Omega$ is transitive, so is $\alpha(\phi)$.

5944. *Proposed by L. J. Wallen, University of Hawaii*

Let H be a separable, complex, infinite-dimensional Hilbert space. A venerable theorem of Halmos states that every contraction is the weak limit of a sequence of unitaries. What is the weak sequential closure of the class of operators similar to unitaries?

5945. *Proposed by R. Sivaramakrishnan, Engineering College, Trichur, India*

Kesava Menon defines the *norm* $f^*(n)$ of a multiplicative function $f(n)$ by

$$f^*(n) = \sum_{d|n^2} f(n^2/d)\lambda(d)f(d),$$

in which $\lambda(n) = (-1)^{k(n)}$ where $k(n)$ represents the total number of prime factors of n , each being counted according to its multiplicity.

Characterize the class of multiplicative arithmetic functions $f(n)$ which satisfy $f^*(n) = [1/n]$, $[x]$ being the integral part of x .

SOLUTIONS OF ADVANCED PROBLEMS

Partitions with Even Minimal Part

5865 [1972, 668]. *Proposed by G. E. Andrews, Pennsylvania State University*

Let Q_n denote the set of partitions of n into distinct non-negative parts with an even number as the smallest part. Let $q_e(n)$ (resp. $q_o(n)$) denote the number of elements of Q_n that have an even number (resp. odd number) of even parts. Prove that

$$q_o(n) - q_e(n) = \begin{cases} 1 & \text{if } n \text{ is a square} \\ 0 & \text{otherwise.} \end{cases}$$

Solution by Allen Stenger, Student, Pennsylvania State University. The generating series for $q_o(n) - q_e(n)$ is

$$\begin{aligned} & \sum_{n=0}^{\infty} x^{2n}(1+x^{2n+1})(1-x^{2n+2})(1+x^{2n+3})\cdots \\ &= \prod_{n=1}^{\infty} \{(1-x^{2n})(1+x^{2n-1})\} \cdot \\ & \quad \left\{ 1 + \frac{x^2}{(1+x)(1-x^2)} + \frac{x^4}{(1+x)(1-x^2)(1+x^3)(1-x^4)} + \cdots \right\}. \end{aligned}$$

The series inside the braces is

$$\frac{1}{2} \left\{ \prod_{n=1}^{\infty} \frac{1}{(1-x^{2n})(1+x^{2n-1})} + \prod_{n=1}^{\infty} (1+x^{2n-1}) \right\}.$$

To prove this, use the known identity

$$(1) \quad 1 + \frac{ax}{1-x} + \frac{a^2x^2}{(1-x)(1-x^2)} + \cdots = \prod_{n=1}^{\infty} \frac{1}{1-ax^n};$$

add the results of putting $a = 1$ and $a = -1$, divide by 2, note that

$$\prod_{n=1}^{\infty} \frac{1}{1+x^n} = \prod_{n=1}^{\infty} \frac{1-x^n}{1-x^{2n}} = \prod_{n=1}^{\infty} (1-x^{2n-1}),$$

and replace x by $-x$. Hence our generating series is

$$\frac{1}{2} + \frac{1}{2} \prod_{n=1}^{\infty} \{(1-x^{2n})(1+x^{2n-1})^2\} = \frac{1}{2} + \frac{1}{2} \sum_{m=-\infty}^{\infty} x^{m^2} = \sum_{m=0}^{\infty} x^{m^2},$$

where the next-to-last equality comes from putting $z = 1$ in Jacobi's identity

$$(2) \quad \prod_{n=1}^{\infty} \{(1-x^{2n})(1+zx^{2n-1})(1+z^{-1}x^{2n-1})\} = \sum_{m=-\infty}^{\infty} z^m x^{m^2}.$$

This is the desired result: the coefficient of x^n is 1 if n is a square, and zero otherwise. (The identities (1) and (2) may be found in Hardy and Wright, *Theory of Numbers*.)

Also solved by L. Carlitz, M. G. Greening (Australia), Phil Tracy, and the proposer.

Real Copositive Quadratic Forms

5867 [1972, 780]. *Proposed by D. E. Daykin, Reading University, England*

Let Q be the real quadratic form

$$\sum_{i=1}^4 \sum_{j=1}^4 a_{ij} x_i x_j \quad \text{with } a_{ij} = a_{ji}.$$

How can we ensure that $Q \geq 0$ whenever all $x_i \geq 0$?

Partial solution by R. D. Leitch, Royal Military College of Science, Shrivenham, England. We shall consider the general case and find a necessary and sufficient condition for the quadratic form

$$Q = \sum_{i,j=1}^n a_{ij}x_ix_j, \quad a_{ij} = a_{ji},$$

to be non-negative when all $x_i \geq 0$. We designate the region where $a_i \geq 0$ as the first quadrant. Since Q is homogeneous we need only consider its values on S^{n-1} : $\{x_i \mid \sum x_i^2 = 1\}$. Let U_n be that portion of S^{n-1} lying in the first quadrant. We shall be considering eigenvectors of various symmetric matrices and shall say that an eigenvector is negative if its associated eigenvalue is negative. We need the following lemma.

LEMMA. *Let Q be the quadratic form $\sum a_{ij}x_ix_j$, and let $A = (a_{ij})$. Then, if D is a closed subset of S^{n-1} , such that Q restricted to the boundary of D is non-negative, Q takes negative values in the interior of D if there is a negative eigenvalue of A in D .*

This is proved by considering the local minima of Q on S^{n-1} , using partial differentiation and Lagrange's multipliers.

We shall need to consider the submatrices of A lying along the main diagonal. Let $A(r_1, \dots, r_k)$ be that submatrix of A formed by the r_1 th, r_2 th, \dots , r_k th, rows and columns of A . Let $Q(r_1, \dots, r_k)$ be the quadratic form defined by $A(r_1, \dots, r_k)$. Observe that $Q(r_1, \dots, r_k)$ is Q restricted to the r_1 th, r_2 th, \dots , r_k th coordinates of (x_1, \dots, x_n) , the other coordinates being zero.

THEOREM. *If Q is non-negative in the first quadrant, then*

1. $a_{ii} \geq 0$, $i = 1, \dots, n$,
2. A and $A(r_1, \dots, r_k)$ have no negative eigenvectors in the first quadrant, for every possible (r_1, \dots, r_k) .

Proof. By induction on n . Let $n = 2$, and consider the quadratic form $Q = ax^2 + 2bxy + cy^2$ and matrix

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix}.$$

Clearly, if either or both of a, c are negative, Q is negative along one or both of the coordinate axes. Applying the lemma with $D = U_2$, we have the theorem when $n = 2$.

Suppose we have the theorem up to $n - 1$, and that the conditions of the theorem hold. Then $Q(i)$ is non-negative in the first quadrant for $i = 1, 2, \dots, n$, and in parti-

cular, Q restricted to the boundary of U_n is non-negative. Applying the lemma gives us the theorem.

Editorial comments. (1) Eric Langford notes that the question has been investigated in J. W. Gaddum, *Linear inequalities and quadratic forms*, Pacific J. Math. 8 (1958), 411–414.

(2) Thomas Markham points out that a characterization of copositive quadratic forms attributed to Garsia and Baumert appears as Theorem 4.1. in the paper, *On classes of copositive matrices*, by R. W. Cottle, G. J. Habetler and C. E. Lemke, *Linear algebra and its applications*, 3 (1970), 295–310.

(3) Milan Lustig (The Technical University, Brno, Czechoslovakia) offers the following sufficient condition for $n = 4$: (i) $a_{ii} \geq 0$; (ii) For all $i, j = 1, 2, 3, 4$, there exist $p_{ij} \geq 0$ such that (a) $p_{ii} = 0$, (b) $\sum_{k=1}^4 p_{ik} = 1$, $i = 1, 2, 3, 4$; (c) $a_{ij} \geq 0$ or $a_{ij}^2 \leq p_{ij}p_{ji}a_{ii}a_{jj}$ for $i \neq j$. For $n = 4$, Lustig offers the following necessary condition: (i) $a_{ii} \geq 0$, (ii) $a_{ij} \geq 0$ or $a_{ij}^2 \leq a_{ii}a_{jj}$ for $i \neq j$.

(4) The proposer also suggests the more general problem in which the conditions $x_i \geq 0$ are replaced by $\sum_{i=1}^n b_{ki}x_i \geq 0$, $k = 1, 2, \dots, (m \leq n)$.

Hereditarily Normal Stone-Čech Compactifiers

5870 [1972, 780]. *Proposed by D. J. Lutzer and F. G. Slaughter, Jr., University of Pittsburgh*

For which discrete spaces D is βD hereditarily normal? (βD denotes the Stone-Čech compactification of D .)

Solution by A. A. Jagers, Technische Hogeschool Twente, Enschede, Netherlands. If D is finite then D is compact and βD coincides with the discrete space D . On the other hand, if D is infinite, βD contains a homeomorphic copy of βN where N is the discrete space of normal numbers, and βN contains a subspace X which is not normal (cf. example 3 on p. 133 of R. Engelking, *Outline of General Topology*, (1968)). Hence βD is hereditarily normal if and only if D is finite.

Also solved by R. Dyckhoff (England), Melvin Henriksen, J. H. Weston, Albert Wilansky, and the proposer.

The Equation $\partial f / \partial x = \partial f / \partial y$

5871 [1972, 780]. *Proposed by P. R. Chernoff, University of California, Berkeley*

Let $f(x, y)$ be a real-valued function of two real variables which is separately differentiable. Assume that $\partial f / \partial x = \partial f / \partial y$ everywhere. Must there be a function g of one variable such that $f(x, y) = g(x + y)$? What if we assume *a priori* that f is jointly continuous?

Solution by K. F. Andersen, University of Alberta. The answer is yes; in fact the domain D of f need not be the entire plane. The conclusion holds provided D has the property that the intersection of D with every line of slope -1 is a connected set. If D is such a set, let $D^* = \{(x, y) : (x, y - x) \in D\}$ and put $h(x, y) = f(x, y - x)$

for $(x, y) \in D^*$. Then, for each fixed y , $\{x: (x, y) \in D^*\}$ is connected, and since

$$\frac{\partial h}{\partial x}(x, y) = f_1(x, y - x) + (-1)f_2(x, y - x) = 0$$

$h(x, y) = g(y)$ is independent of x . Hence, $f(x, y) = h(x, x + y) = g(x + y)$ for all $(x, y) \in D$.

Generators for a Function Ring

5873 [1972, 913]. *Proposed by Helge Tverberg, University of Bergen, Norway*

Those real polynomials in x and the greatest integer function $[x]$ which are continuous functions of x form a ring A , containing R . Find the minimal set of generators, over R , of A .

Solution by D. Ž. Djoković, University of Waterloo. Let $f(x) = x$, $g(x) = [x]$ for $x \in R$ and $h = f - g - \frac{1}{2}$. The functions h^2 and $h(h^2 - \frac{1}{4})$ are continuous and if P is a polynomial in two variables we can write

$$\begin{aligned} P(f, g) &= P(f, f - h - \tfrac{1}{2}) = Q(f, h) = Q_1(f, h^2) + hQ_2(f, h^2) \\ &= Q_1(f, h^2) + h(h^2 - \tfrac{1}{4})Q_3(f, h^2) + hQ_4(f, h^2) \end{aligned}$$

where Q_1, Q_2, Q_3, Q_4 are suitable polynomials. It follows that $P(f, g)$ is continuous if and only if $Q_4 = 0$. Thus A is generated over R by the functions

$$f, h^2 \quad \text{and} \quad h(h^2 - \tfrac{1}{4}).$$

We claim that A cannot be generated by two elements over R . If this were so then A would be isomorphic to the polynomial algebra $R[X, Y]$ since f and h^2 are algebraically independent over R and A is an integral domain. Thus A would be a unique factorization domain but if $y = h^2$, $z = h(h^2 - \frac{1}{4})$ then

$$(1) \quad z^2 = y(y - \tfrac{1}{4})^2.$$

We have $A = R[f, y, z]$ and one can see easily that the elements $z, y, y - \frac{1}{4}$ are irreducible in A . Indeed, $A \subset R[f, h]$ and every factorization in A would give a factorization in $R[f, h]$ and these are all known for the elements $z, y, y - \frac{1}{4}$. Then (1) shows that A is not a unique factorization domain, which is a contradiction.

Also solved by the proposer and I. Beck (Norway).

INDEX TO VOLUME 80, 1973 THE AMERICAN MATHEMATICAL MONTHLY

Author Index	1187
Key Words and Phrases Index	1190
Problems and Solutions Index	1194
Reviews Index	1196
News and Notices Index	1217
MAA and its Sections Index	1218

AUTHOR INDEX

- AICHELE DB Training secondary mathematics teachers in Venezuela 798–803
- ALAS OT On set points of discontinuity 186–187
- ALTER RONALD Can $\phi(n)$ properly divide $n-1$? 192–193
- APOSTOL TM Another elementary proof of Euler's formula for $\zeta(2n)$ 425–431
- ASSOCIATION FOR WOMEN IN MATH Remarks on "Women in mathematics" 903–904
- AULT JC AND WATTERS JF Circle groups of nilpotent rings 48–52
- Award for Distinguished Service to Professor Raymond L Wilder 117–119
- Award of the 1973 Chauvenet Prize to Professor Carl D Olds 120
- BARKER GP Topological properties of the row echelon form 787–789
- BARNES CW Remarks on the Bessel polynomials 1034–1040
- BENDER EA Teaching applicable mathematics 302–307
- BENDER EA AND NEUWIRTH LP Traffic flow: Laplace transforms 417–423
- BILLINGSLEY PATRICK Prime numbers and Brownian motion 1099–1115
- BIRKHOFF GARRETT Current trends in algebra 760–782
- BIRNBAUM S The main crises 545–546
- BOAS RP and POLLARD H Continuous analogues of series 18–25
- BOLKER ED The spinor spanner 977–984
- BRAWLEY JV and CARLITZ L A characterization of the $n \times n$ matrices over a finite field 670–672
- and ——— An addendum to the paper "A characterization of the $n \times n$ matrices over a finite field" 1041–1043
- BROWNE JB See Eisenberg TA
- BRUCKNER AM The differentiability properties of typical functions in $C[a, b]$ 679–683
- BURROW MD The minimal polynomial of a linear transformation 1129–1131
- CALLAHAN FP An identity satisfied by derivations of a purely inseparable field 40–42
- CAMERON DE The mini-max property of the Tychonoff product topology 925–927
- CARLITZ L and SCOVILLE RICHARD The sign of the Bernoulli and Euler numbers 548–549
- CARLITZ L Inequalities for the area of two triangles 910–911
- CARLITZ L See Brawley JV
- CARROLL JJ FISHER GA ODLYZKO AM SLOANE NJA What are the Latin square groups? 1045
- CHAKERIAN GD and KLAMKIN MS Inequalities for sums of distances 1009–1017
- CHINN WG See Gilmer GF
- CLEVELAND RICHARD A global characterization of uniform continuity 64–66
- COHN PM Unique factorization domains 1–18
- Correction to "Unique factorization domains" 1115
- COLLINS GE Computer algebra of polynomials and rational functions 725–755
- COOKE WP Geometric fit of a monotonic cubic 1047–1051
- CUPM Report to the Board of Governors August 1972 313–314
- DARST RB Simple proofs of two estimates for e 194
- DAVIS MARTIN Hilbert's tenth problem is unsolvable 233–269
- DEAKIN MAB Developing countries: A rejoinder 806

- DEAVOURS CA The quaternion calculus 995–1008
- DE BOER DELMER See Williams HE
- DERRICK WR A condition under which mapping is a homeomorphism 554–555
- DONNELLY HAROLD On a problem concerning Euler's ϕ -function 1029–1031
- DRIVER RD SASSER DW and SLATER ML The equation $x'(t) = ax(t) + bx(t - \tau)$ with "small" delay 990–995
- DROBOT VLADIMIR On sums of powers of a number 42–44
- EGGAN LC and INSEL AJ A Wronskian condition related to ordinary differential equations 300–302
- EIDSWICK JA A crowded set of nonintersecting lines 415
- EISENBERG TA and BROWNE JB Using student-tutors in precalculus instruction 685–688
- ELLIS DF Economics as a minor for undergraduate mathematics majors 688–689
- ERDÖS P and GUY RK Crossing number problems 52–58
- FADELL AG A proof of the chain rule for derivatives in N -space 1134–1135
- FISHER B On a problem of Besicovitch 785–787
- FISHER GA See Carroll JJ
- FLANDERS HARLEY Differentiation under the integral sign 615–627
- , Number fifty-two 1099
- FLANIGAN FJ Some half-plane Dirichlet problems: A bare hands approach 59–61
- FRELICH GERALD Increasing continuous singular functions 918–919
- FUSARO BA The area of a hypersphere in Riemannian space 179–184
- GALE DAVID On the theory of interest 853–868
- GILMER R and ROSELLE D Complements and comments 1116–1118
- GILMER GF SILER HB MANSFIELD R and CHINN WG Concerns of two year colleges 1055–1057
- GOLDSTEIN LJ A history of the prime number theorem 599–615
- , Correction to "A history of the prime number theorem" 1115
- GORDON WB Addendum to "On the diffeomorphisms of Euclidean space" 674–675
- GRAUDONS NANCY See Parberry EA
- GREENSPAN HP Applied mathematics at M. I. T. 67–72
- GREENSPAN DONALD A finite difference proof that $E = mc^2$ 289–292
- GREITZER S The first USA Mathematical Olympiad 276–281
- GUSTAFSON WH What is the probability that two group elements commute? 1031–1034
- GUY RICHARD Monthly Research Problems 1969–1973 1120–1128
- GUY RK See Erdős P
- GUY RK and SELFRIDGE JL The nesting and roosting habits of the laddered parenthesis 868–876
- HAHN LIANG-SHIN On an extension of the theorem of Hausdorff-Young 667–669
- HALMOS PR The legend of John von Neumann 382–394
- HAM MW The lecture method in mathematics: A student's view 195–201
- HANNA AZMI On injective modules 297–298
- HEIMER RT See Jansson LC
- HEINEN JA and WILANSKY ALBERT A theorem on set inclusion in metric spaces 46–48
- HENRY BR Solution of Fejes Tóth's Illumination Problem 409–410
- HERSH RFUBEN How to classify differential polynomials 641–654
- HIGGINS JJ Representing a finite Borel measure in terms of its distribution function 683–685
- HINDMAN NEIL Basically bounded sets and a generalized Heine-Borel theorem 549–552
- HIRSCHHORN MD How unexpected is the prime number theorem? 675–677
- HORADAM AE See Shannon AG
- HUGHS BB Survival for mathematics students 689–690
- INSEL AJ See Eggan LC
- JACOB HG Another proof of the rational decomposition theorem 1131–1134
- JAMESON GJO A problem on series 1119
- JANSSON LC and HEIMER RT On behavioral objectives in mathematics education 930–933
- JENNER WE On non-associative algebras derived from graphs 288–289
- JONES JP and TOPOROWSKI S Irrational numbers 423–424
- KANTER MAREK Stable laws and the imbedding of L^p spaces 403–407
- KAZARINOFF ND and WEITZENKAMP ROGER Squaring rectangles and squares 877–888

- KEMENY JG What every college president should know about mathematics 889–901
- KIEFFER JC A covering theorem 410–411
- KILPATRICK J See Pólya G
- KIMBERLING CH Two-dimensional complete monotonicity with diagonalization 789–791
- KLAMKIN MS See Chakerian GD
- KLARNER DA and RADO R Linear combinations of sets of consecutive integers 985–989
- LACEY HE The Hamel dimension of any infinite dimensional separable Banach space is c 298
- LANGFORD ERIC Distributivity over the Dirichlet product and completely multiplicative arithmetical functions 411–414
- LARNEY VIOLET H Female mathematicians, where are you? 310–313
- LEIBOWITZ GERALD The Cesaro operators and their generalizations: Examples in infinite-dimensional linear analysis 654–661
- LEONARD JL A discovery course in graph theory 1052–1053
- LINDSTROM PA Survival of the two-year college mathematics teacher 1135–1137
- LUXEMBURG WAJ What is nonstandard analysis? part II 38–67
- MANSFIELD R See Gilmer GF
- McKAY JH The William Lowell Putnam Mathematical Competition 170–179, 1017–1028
- McSHANE EJ A unified theory of integration 349–359
- The Lagrange multiplier rule 922–925
- MERRIS RUSSELL The permanent of a doubly stochastic matrix 791–793
- MEYER WALTER Equitable coloring 920–922
- MILLMAN RS and STEHNEY ANN K The geometry of connections 475–500
- MINASSIAN DP Types of fully ordered groups 159–169
- MINSKER STEVEN A familiar combinatorial identity proved by complex analysis 1051
- MONNA AF Experiences with lectures on the history of mathematics in Utrecht 803–806
- MOORE MH A convex matrix function 408–409
- MORDELL LJ The sign of the Bernoulli numbers 547–548
- MOSER LEO Some mathematical verses 902
- NADLER SB The indecomposability of the dyadic solenoid 677–679
- NASH BO Reachability problems in vector addition systems 292–295
- NEUWIRTH LP See Bender EA
- NICKEL PA Single layer potentials and the Cauchy-Kowalewski theorem 61–64
- ODLYZKO AM On lattice points inside convex bodies 915–918
- See Carroll JJ
- O'HARA PJ Another proof of Bernstein's theorem 673–674
- PAPADIMITRIOU IOANNIS A simple proof of the formula $\sum_{k=1}^{\infty} k^{-2} = \pi^2/6$ 424–425
- PAPANICOLAOU GC Stochastic equations and their applications 526–545
- PARBERRY EA and GRAUDONS NANCY When do all k -sequences modulo m have period one? 295–297
- PARKINSON CLAIRE Ambivalence in alternating symmetric groups 190–192
- PIZER AK A problem on rational functions 552–553
- POLLARD H See Boas RP
- PÓLYA G A letter by professor Pólya 73–74
- PÓLYA G and KILPATRICK J The Stanford University competitive examination in mathematics 627–640
- PUTNAM HILARY Recursive functions and hierarchies, part II 68–86
- RADO R See Klarner DA
- RAMALEY WC Independent study for undergraduates 555–558
- RANDES RH and SCHAEFFER AJ An integrated sequence in the mathematical sciences for undergraduate business students 431–433
- RECAMÁN S BERNARDO Questions on a sequence of Ulam 919–920
- RIORDAN JOHN A note on Catalan parentheses 904–906
- ROBINSON ABRAHAM Function theory on some nonarchimedean fields, part II 87–109
- ROSELLE D See Gilmer R
- ROSMAN BH Another approach to the cubic interpolating spline 927–930
- RUCHTE MF and RYDEN RW A proof of uniqueness of factorization in the gaussian integers 58–59
- RUDIN WALTER A generalization of a theorem of Archimedes 794–796
- RYDEN RW See Ruchte MF
- SAMUELSSON ÅKE A local mean value theorem for analytic functions 45–46
- SASSER DW See Driver RD

- SCHAEFFER AJ See Randles RH
 SCHOENBERG IJ The elementary cases of Landau's problem of inequalities between derivatives 121-158
 SCOTT EJ Determination of the Riemann function 906-909
 SCOVILLE RICHARD See Carlitz L
 SELFRIDGE JL See Guy RK
 SHANNON AG and HORADAM AF Generalized Fibonacci number triples 187-190
 SHAPIRO HN A micronote on a functional equation 1041
 SHISHA O On the discrete version of Wirtinger's inequality 755-760
 SIELAFF RW Perfect parallelograms 414-415
 SILVER JERRY and WAITS BERT Multiple-choice examinations in mathematics not valid for everyone 937-942
 Simple groups 1028
 SINNER HB See Gilmer GF
 SLATER ML See Driver RD
 SLOANE NJA See Carroll JJ
 SPENCER JOEL A deception game 416-417
 STANFORD DP Functions satisfying a mean value property at their zeros 665-667
 STEEN LA Highlights in the history of spectral theory 359-381
 STEHNEY ANN K See Millman RS
 SWENSON JR The chromatic polynomial of a complete bipartite graph 797-798
 TAYLOR PD A Banach space characterization of the space of affine continuous functions on a compact convex set 911-915
 TOPOROWSKI S See Jones JP
 TÓTH FEJES L Exploring a planet 1043-1044
 TULL JP A discovery approach to e 193-194
 ULLMAN JL An area theorem for Schlicht functions 184-186
 VAN VLECK FS A remark concerning absolutely continuous functions 286-287
 VAUGHT RL Some aspects of the theory of models, part II 3-37
 WAITS BERT Individualized instruction in large enrollment mathematics courses 307-310
 ——— See Silver Jerry
 WALSH JL History of the Riemann mapping theorem 270-276
 WALTER JOHANN On elementary proofs of Peano's existence theorems 282-286
 WATERS JF See Ault JC
 WEGNER BERND Existence of four concurrent normals to a smooth closed hypersurface of E^n 782-785
 WEITZENKAMP ROGER See Kazarinoff ND
 WILANSKY ALBERT See Heinen JA
 WILLCOX AB England was lost on the playing fields of Eton: A parable for mathematics 25-40
 WILLIAMS HE and DE BOER DELMER Teaching a computer-oriented laboratory course for ordinary differential equation 933-937
 WILLIAMS RK A note on conformality 299-300
 WILLMORE T Correction to "The Math Societies and Associations in the U.K." 876
 WILSON RJ An introduction to matroid theory 500-525
 WILSON RL A bow to relevancy 1053-1055
 WYMAN BF Correction to: What is a reciprocity law? 281
 ZAHN CT Alternating Euler paths for packings and covers 395-403
 ZELINSKY D A.A. Albert 661-665

KEY WORDS AND PHRASES INDEX

- Absolutely continuous functions VAN VLECK FS 286
 Additive functions SHAPIRO HN 1041
 Albert AA ZELINSKY D 661
 Algebra BIRKHOFF G 760
 Algorithms COLLINS GE 725
 Alternating groups PARKINSON C 190
 Analytic function WILLIAMS RK 299
 Applicable mathematics BENDER EA 302
 Applications WILSON RL 1053
 Applied mathematics education GREENSPAN HP 67
 Approximation by power sums DROBOT V 42
 Area theorem ULLMAN JL 184
 Awards 117 120
 Banach space HEINEN JA & WILANSKY A 46
 Behavioral objectives JANSSON LC & HEIMER RT 930

- Bernoulli numbers APOSTOL TM 425 MORDELL LJ 547 CARLITZ L & SCOVILLE R 548
 Bernstein's theorem O'HARA PJ 673
 Bessel polynomials BARNES CW 1034
 Bipartite graph SWENSON JR 797
 Borel measure HIGGINS JJ 683
 Brownian motion BILLINGSLEY P 1099
 Business students RANDLES RH & SCHAEFFER AJ 431

 Calculus course BENDER EA 302
 Catalan numbers GUY RK & SELFRIDGE JL 868 RIORDAN J 904
 Cauchy problem SCOTT EJ 906
 Cesaro operators LEIBOWITZ G 654
 Chain rule FADELL AG 1134
 Characteristic polynomial BURROW MD 1129
 Chromatic number MEYER W 920
 Circle group AULT JC & WATTERS JF 48
 College president KEMENY JG 889
 College teaching HUGHES BB 689
 Coloring MEYER W 920
 Combinatorial identity MINSKER S 1051
 Committee on two-year colleges GILMER GF SINNER HB MANSFIELD R & CHINN WG 1055
 Commuting elements group GUSTAFSON WH 1031
 Compactness theorem VAUGHT RL June-July part II 3
 Competition Putnam McKAY JH 170 1017
 Complements and comments GILMER R & ROSELLE DAVID 1116
 Completely monotone matrices KIMBERLING CH 789
 Computer algebra COLLINS GE 725
 Computer differential equations WILLIAMS HE & DE BOER D 933
 Concurrent normals WEGNER B 782
 Conformal mapping WALSH JL 270 WILLIAMS RK 299
 Connections MILLMAN RS & STEHNEY AK 475
 Consecutive integers KLARNER DA & RADO R 985
 Convex sets TAYLOR PD 911 FEJES TÓTH L 1043
 Covering theorem KIEFFER JC 410
 Crises in mathematics BIRNBAUM S 545
 Crossing numbers ERDÖS P & GUY RK 52 CUPM 313

 Decay equation DRIVER RD SASSER DW & SLATER ML 990
 Deception game SPENCER J 416
 Decision making KEMENY JG 889
 Derivations CALLAHAN FP 40
 Developing countries DEAKIN MAB 806
 Diffeomorphisms GORDON WB 674
 Differentiability BRUCKNER AM 679
 Differential-difference equation DRIVER RD SASSER DW & SLATER ML 990
 Differential geometry MILLMAN RS & STEHNEY AK 475
 Differential polynomials HERSH R 641
 Differentiation under the integral sign FLANDERS H 615
 Diophantine equation DAVIS M 233
 Dirichlet problems FLANIGAN FJ 59
 Discovery course LEONARD JL 1052
 Distribution function COOKE WP 1047
 Doubly stochastic matrix MERRIS R 791

 e TULL JP 193 DARST RB 194
 Economics GALE D 853
 Economics minor ELLIS DF 688
 Electric circuits KAZARINOFF ND & WEITZENKAMP R 877
 Energy and mass GREENSPAN D 289
 Euler function DONNELLY H 1029
 Euler numbers CARLITZ L & SCOVILLE R 548
 Euler paths ZAHN CT 395
 Euler's formula PAPADIMITRIOU I 424 APOSTOL TM 425
 Euler totient ALTER R 192

 Fibonacci triples SHANNON AG & HORADAM AF 187

 Gaussian integers RUCHTE MF & RYDEN RW 58
 Geometric inequalities CHAKERIAN GD & KLAMKIN MS 1009
 Graphs JENNER WE 288
 Graphs in the plane ERDÖS P & GUY RK 52
 Graph theory WILSON RJ 500 LEONARD JL 1052
 Group, commuting elements GUSTAFSON WH 1031

 Hamel dimension LACEY HE 298
 Harmonic space FUSARO BA 179
 Hausdorff-Young theorem HAHN LS 667

- Heine-Borel theorem HINDMAN N 549
 Hierarchies PUTNAM H June-July part II 68
 Hilbert space STEEN LA 359
 Hilbert's 10th problem DAVIS M 233
 History of algebra BIRKHOFF G 760
 History of mathematics MONNA AF 803
 Homeomorphism DERRICK WR 554
 Hyperbolic differential equation SCOTT EJ 906
 Hypersphere FUSARO AB 179

 Illumination problem HENRY BR 409
 Indecomposable continua NADLER SB 677
 Independent study RAMALEY WC 555
 Individual instruction WAITS B 307
 Inequalities CHAKERIAN GD & KLAMKIN MS 1009
 Inequalities between derivatives SCHOENBERG IJ 121
 Infinite integrals BOAS RP & POLLARD H 18
 Infinite series BOAS RP & POLLARD H 18 JAMESON GJO 1119
 Infinitesimals LUXEBURG WAJ June-July part II 38
 Injective module HANNA A 297
 Integration McSHANE EJ 349
 Interest GALE D 853
 Irrational numbers JONES JP & TOPOROWSKI S 423

 Makeya problem FISHER B 785
 Knots BOLKER ED 977

 Lagrange multipliers McSHANE EJ 922
 Landau's problem SCHOENBERG IJ 121
 Laplace transforms BENDER EA & NEUWIRTH LP 417 SCOTT EJ 906
 Large courses WAITS B 307
 Latin squares CARROLL JJ FISHER GA ODLYZKO AM SLOANE N 1045
 Lattice points ODLYZKO AM 915
 Lebesgue integral McSHANE EJ 349
 Lecture method HAM MW 195
 Leibnitz rule FLANDERS H 615
 Linear connections MILLMAN RS & STEHNEY AK 475
 Lowenheim-Skolem theorem VAUGHT RL June-July part II 3
 L^p spaces KANTER M 403

 Markov processes PAPANICOLAOU GC 526
 Mathematics Association WILLMORE T 876

 Mathematics contest PÓLYA G & KILPATRICK J 627
 Matrices over a finite field BRAWLEY JV & CARLITZ L 670 1041
 Matrix function MOORE MH 408
 Matroid WILSON RJ 500
 Mean value property STANFORD DP 665
 Mean value theorem SAMUELSSON Å 45
 Metric spaces HEINEN JA & WILANSKY A 46
 Minimal polynomial BURROW MD 1129
 Models VAUGHT RL June-July part II 3 KEMENY JG 889
 Moser Leo 902
 Multiple choice examinations SILVER J & WAITS B 937
 Multiple functions LANGFORD E 411

 Nilpotent ring AULT JC & WATTERS JF 48
 Non-archimedean fields ROBINSON A June-July part II 87
 Non-intersecting lines EIDSWICK JA 415
 Non-standard analysis ROBINSON A 87 VAUGHT RL June-July part II 3
 Normals of a hypersurface WEGNER B 782
 Numerical algebra BIRKHOFF G 760

 Olympiad GREITZER SL 276
 Ordered fields ROBINSON A June-July part II 87
 Ordered groups MINASSIAN DP 159

 Parable WILLCOX AB 25
 Parentheses GUY RK & SELFRIDGE JL 868 RIORDAN J 904
 Partially ordered groups MINASSIAN DP 159
 Peano's existence theorem WALTER J 282
 Perfect parallelograms SIELAFF RW 414
 Permanent MERRIS R 791
 Plank problem FEJES TÓTH L 1043
 Polynomials COLLINS GE 725
 Precalculus instruction EISENBERG TA & BROWNE JB 685
 Prime number BILLINGSLEY P 1099
 Prime number theorem GOLDSTEIN LJ 599 HIRSCHHORN MD 675
 Problem solving PÓLYA G 73
 Product topology CAMERON D 925
 Promotion PÓLYA G 73
 Putnam competition McKAY JH 170 1017

 Quaternion calculus DEAVOURS CA 995

- Rational canonical form JACOB HG 1131
 Rational functions PIZER AK 552
 Reciprocity law WYMAN BF 281
 Recursive functions PUTNAM H June-July part II 68
 Relevance WILLCOX AB 25
 Research problems GUY RK 1120
 Riemann mapping theorem WALSH JL 270
 Row echelon form BARKER GP 787

 Schlicht functions ULLMAN JL 184
 Sequences RECAMÁN B 919
 Sequences of integers PARBERRY EA & GRAUDONS N 295
 Set of discontinuity ALAS OT 186
 Simple groups 1028
 Single layers NICKEL PA 61
 Singular functions FREILICH G 918
 Space volume RUDIN W 794
 Spectral theory STEEN LA 359
 Spinor BOLKER ED 977
 Splines SCHOENBERG IJ 121 ROSMAN BH 927
 Squaring rectangles KAZARINOFF ND & WEITZENKAMP R 877
 Stanford competition PÓLYA G & KILPATRICK J 627
 Stochastic equations PAPANICOLAOU GC 526
 Stochastic integral KANTER M 403
 Student tutors EISENBERG TA & BROWNE JB 685

 Summability theory LEIBOWITZ G 654
 Summation of \sum^{n-2} PAPADIMITRIOU I 424
 Survival LINDSTRÖM PA 1135

 Teacher training AICHELE DB 798
 Teaching HUGHES BB 689
 Totient DONNELLY H 1029
 Traffic flow BENDER AE & NEUWIRTH LP 417
 Triangles CARLITZ L 910
 Turing machines PUTNAM H June-July part II 68
 Two-year colleges GILMER GF SINER HB MANSFIELD R & CHINN WG 1055
 Two-year college teacher LINDSTRÖM PA 1135

 Uniform continuity CLEVELAND R 64
 Unique factorization RUCHTE MF & RYDEN RM 58
 Unique factorization domain COHN PM 1 1115

 Vector addition system NASH BO 292
 Venezuela AICHELE DB 798
 Von Neumann John HALMOS PR 382

 Wirtinger's inequality SHISHA O 755
 Women in mathematics LARNEY VH 310 ASSOC. FOR WOMEN IN MATH. 903
 Wronskian EGGAN LC & INSEL AJ 300

 Zeta function APOSTOL TM 425

PROBLEMS AND SOLUTIONS

PROBLEMS PROPOSED

- | | | |
|---------------------------|---------------------------|-------------------------------|
| Alexander JC 440 | Grosch CB 435 | O'Brien George 943 |
| Al Salam WA 315 | Hahn L.-S 943 | O'Farrell AG 814 |
| Andrushkiw JW 1067 | Harris LA 697 | Pomerance Carl 949 |
| Apostol TM 1058 | Hemperly JC 1058 | Recamán Bernardo 434 1057 |
| Barnes FW 434 | Heuer GA 325 | Redheffer Ray 1138 |
| Battany DM 565 | Hoshek Lyles 691 | Reese Sylvester 209 |
| Bernhart Frank 208, 324 | Howard FT 559 | Reingold EM 691 |
| Bloom DM 1147 | Hsieh SC 325 | Renz PL 949 |
| Boas RP 814 | Hubbard JH 324 | Ringel CM 82 |
| Boyd AV 434 | Hwang FK 1058 | Ruderman HD 82 807 1139 |
| Brons KA 202 | Ivanoff VF 203 | Schaumberger Norman 316 |
| Buck RC 691 1067 | Jackson DE 1058 | Schurle AW 209 |
| Buckley JJ 949 | Johnson CR 814 | Selucky K 1067 |
| Chen SY 325 | Johnson Wells 943 | Shafer RE 316 |
| Cohn Paul 697, 814 | Just Erwin 76 316 | Singmaster David 75 |
| Cooper CDH 559 | Kamp JF 83 | Sivamakrishnan R 1147 |
| Dashiell FK 83 | Kirk RB 564 | Siwec Frank 1139 |
| Daykin DE 202 564 | Klamkin MS 75 75 807 | Smythe RT 943 |
| Deutsch Emeric 814 | Knight Bill 564 | Smyth CJ 949 |
| Dixon ED 324 | Kuzam FA 82 | Spencer Joel 209 |
| Dlab V 82 | Laugwitz Detlef 1067 | Stanley Richard 949 |
| Dodge CW 202 | Letac Gerard 440 441 1139 | Stern Frederick 949 |
| Doyle JK 82 | Leuenberger F 1138 | Stewart BM 691 |
| Dugdale JK 564 | Linderholm CE 564 | Styer David 1067 |
| Eakin PM 441 | Long CA 807 | Tomescu Ioan 559 |
| Ehrenfeucht Andrzej 697 | Masley John 808 | Umberger Edmund 1058 |
| Entringer RC 1058 | Mauldon JG 697 | Uppuluri VRR 74 |
| Gentile ER 324 | Maurer Russell 316 | Walker AW 202 316 560 |
| Girod Donald 814 | McConnell Alan 808 | Wallen LJ 1147 |
| Glasser ML 440 | McLean David 943 | Wang ETH 691 692 808 943 1139 |
| Goldberg Michael 434 | Minassian Donald 1146 | Washington LC 697 |
| Golomb SW 697 | Murray PJ 692 | Wendel JG 325 559 |
| Good IJ 209 | Mycielski Jan 1146 | Wilansky Albert 325 564 1067 |
| Gould HW 75 | Myhill John 83 | Wolk Barry 434 |
| Greitzer SL 75 | Newman David 203 | |
| Groenewoud Cornelius 1058 | Nicol CA 560 692 | |

PROBLEMS SOLVED

- | | | |
|--------------------------------------------|-------------------------|----------------------------------|
| Andersen KF 1150 | Carty Frederick 317 813 | Gardner CS 1139 |
| Annulis JT 1071 | Chakerian GD 562 | Garfield Ralph 321 |
| Atkinson DW 1144 | Charnow Allen 1061 | Gerber Leon 1145 |
| Bauman Norman 561 | Chouteau Charles 693 | Gerst Irving 214 |
| Belanger DG 567 1068 | Comiskey John 320 | Gibbs Richard 1066 |
| Bennett Coll. Team 1060 | Converse GA 562 | Gilmer Robert 944 |
| Bern Switzerland Problem Solving Group 319 | Coolidge John 947 | Goldberg Michael 692 |
| Bernau SJ 87 | Coppersmith Don 815 | Greening MG 436 439 694 695 1143 |
| Bloom DM 206 320 563 1142 | D'Alarcao H 320 | Grimm CA 80 |
| Borwein David 698 | Davies RO 87 | Grossman JW 210 |
| Breiteig Trygve 696 | DeMeyer Frank 83 | Gudder SP 566 |
| Breusch Robert 809 | Dickson RJ 435 438 | Guggenheimer H 211 |
| Buck RC 1144 | Djokovic DZ 1151 | Hertz Ellen 811 951 |
| Burke PJ 207 | Evans RJ 560 | Heuer GA 952 |
| Buschman RG 213 | Felsinger Neal 327 | Hobbs AM 950 |
| Butler JG 443 | Ferrero Bruce 1069 | Huddleston Nancy 700 |
| Carlitz Leonard 441 819 | Fine NJ 435 | Israel RB 329 |
| | Galvin Fred 950 | Jagers AA 85 815 1150 |

Johnson Wells 207	Moore T 320	Stein Alan 1059
Jondrup Soren 212	Moser WOJ 437	Stenger Allen 1148
Kappus Hans 810	Niven Ivan 812	Stocker Harold 1066
Kestelman H 1062	Pierce Stephen 443	Stoll Manfred 698
Klamkin MS 323 1144	Prielipp Bob 439	Tang Hwa 1061
Klasi ML 80	Prostanstus LP 438	Taylor Herbert 695
Klein EM 326	Quackenbush RW 950	Taylor RL 698
Kuenzi NJ 439	Robinson GB 77 203	Temple Univ Probl Solv Group 948
Langford ES 1063	Rose-Hulman Sol Group 1143	Thomas Gomer 818
Lass Harry 84 1070	Rosenthal Eric 445	Thomas Martin 808
Leibowitz GM 566 568	Rousseau CC 328 700	Torchinelli Guy 1064
Leitch RD 1149	Ruehr OG 1071	Trost EW 77
Linus V 1144	St Olaf College Students 436 945 1146	Vitale Mike 946
Little JG 1071	Schmitt FG 701	Walker AW 203 204 205
Lossers OP 818 1064	Schuurmann Fred 817	Wall CR 696
Makowski Andrzej 78	Scoville Richard 79 441	Wetzel JE 562
Mattics LE 1139	Segal AC 83	Wilkins JE Jr 1061
Mauldon JG 79 81 946	Smith JR 1068	Wilson Norman 1062
Meir Amram 444	Snow Wolfe 565	Witsenhausen HS 318
Mitra SS 206	Spear David 86	Wong William 442
Monk David 317	Spindler Stephen 76	Editorial Note 561
Montgomery PL 446	Starke EP 77	

SOLUTIONS

Numbers in **boldface** type refer to problems, those in lightface, to pages

E-1085 808	E-2245 809	E-2293 317	E-2294 692	5814 83	5815 83	5816 84	5817 85
E-2330 76	E-2332 77	E-2333 78			5818 86	5820 87	5821 210
E-2334 79	E-2335 80	E-2336 81			5822 211	5823 212	5824 213
E-2337 203	E-2338 204	E-2339 205			5825 214	5826 326	5827 326*
E-2340 206	E-2341 206	E-2342 207			5828 327	5829 327	5830 328
E-2343 317	E-2345 318	E-2346 319			5831 329	5832 441	5833 442
E-2347 321	E-2348 323	E-2350 435			5834 443	5835 443	5836 444
E-2351 436	E-2352 436	E-2353 437			5837 445	5838 445	5839 446
E-2354 439	E-2355 560	E-2356 810			5840 565	5841 566	5842 566
E-2357 439	E-2358 561	E-2359 561			5843 567	5844 568	5845 698
E-2360 562	E-2361 693	E-2362 810			5846 698	5847 699	5848 700
E-2363 694	E-2364 563	E-2365 695			5849 701	5850 815	5851 815
E-2366 695	E-2367 696	E-2368 812			5852 817	5853 818	5854 819
E-2369 813	E-2370 944	E-2371 945			5855 950	5856 950	5857 951
E-2372 946	E-2373 1139	E-2374 946			5858 952	5859 1068	5862 1068
E-2375 947	E-2376 948	E-2378 1059			5863 1069	5864 1070	5865 1147
E-2379 1060	E-2380 1061	E-2381 1062			5867 1148	5868 1070	5869 1071
E-2382 1064	E-2383 1066	E-2386 1141			5870 1150	5871 1150	5873 1151
E-2388 1142	E-2391 1143	E-2393 1145					
E-2395 1146							

* Editorial note 326

REVIEWS

Names of authors are in ordinary type, those of reviewers in capitals.

- Apostol TM *Selected Papers on Calculus* DOROTHY K BERNSTEIN 93-94
- Bechtell Homer *The Theory of Groups* RE PHILLIPS 447
- Beck Anatole Bleicher MN and Crowe DW *Excursions into Mathematics* JAMES STASHEFF 821
- Behrens Ernest-August *Ring Theory* N DIVINSKY 95
- Blakeslee DW and Chinn WG *Introductory Statistics and Probability: A Basis for Decision Making* DS MOORE 214
- Bleicher MN See Beck Anatole
- Blumenthal LM and Menger Karl *Studies in Geometry* A BRUEN 330
- Brauer Fred Nohel JA Schneider Hans *Linear Mathematics: An Introduction to Linear Algebra and Linear Differential Equations* WS LOUD 451
- Campbell HG *Linear Algebra with Applications: Including Linear Programming* DE CHRISTIE 702
- Chinn WG See Blakeslee DW
- Chover Joshua *The Green Book of Calculus* AD KRAMER 955-956
- Cohn PM *Free Rings and their Relations* DJ FIELDHOUSE 573
- Crowe DW See Beck Anatole
- Davenport WB JR *Probability and Random Processes* JL SNELL 88-90
- Durbin JR *Mathematics: Its Spirit and Evolution* RANDALL LONGCORE 1074-1075
- Dwass Meyer *Probability Theory and Applications* JL SNELL 88-90
- Edwards AL *Probability and Statistics* DS MOORE 215
- Embry Mary Schell JF Thomas JP *Calculus and Linear Algebra: an Integrated Approach* RT HOOD 454
- Ericksen GL *Scientific Inquiry in the Behavioral Sciences: an Introduction to Statistics* DS MOORE 215
- Finkbeiner DT *Elements of Linear Algebra* DE CHRISTIE 702
- Folks Leroy See Kempthorne Oscar
- Gillett Philip *Linear Mathematics* WS LOUD 451
- Goodman AW Ratti JS *Finite Mathematics with Applications* ELIZABETH BERMAN 91-92
- TW CASSTEVENS 91
- Gratzner George *Lattice Theory: First Concepts and Distributive Lattices* DT FINKBEINER 824
- Gray Mary W *Calculus with Finite Mathematics for Social Sciences* GERALD GIACCAI and KENNETH SLONNIGER 1076-1077
- Herstein IN and Sandler R *Introduction to the Calculus* JV LEWIS 90-91
- Hilton PJ and Stambach U *A Course in Homological Algebra* CLAUDE SCHOCHET 1153-1154
- Kaplansky Irving *Set Theory and Metric Spaces* WR PARK 953-955
- Kempthorne Oscar and Folks Leroy *Probability, Statistics and Data Analysis* J KIEFER 822
- Kochendorffer Rudolf *Group Theory* RE PHILLIPS 447
- Kreyszig Erwin *Introductory Mathematical Statistics* BP KORIN 454
- Larsen MD McCarthy PJ *Multiplicative Theory of Ideals* HH BRUNGS 94
- LeLionnais F (ed.) *Great Currents of Mathematical Thought* KO MAY and SB REGO-CZEI 825
- Levine Arnold *Theory of Probability* LJ SNELL 88-90
- Long PE *An Introduction to General Topology* HELEN E SALZBERG 1077
- Lukacs Eugene *Probability and Mathematical Statistics: an Introduction* DS MOORE 214
- Maunder CRF *Algebraic Topology* MJ POWERS 449
- McCarthy PJ See Larsen MD
- Mendenhall W *Introduction to Probability and Statistics, Third Edition* DS MOORE 215
- Menger Karl See Blumenthal LM
- Meyer PL *Introductory Probability and Statistical Applications 2nd edition* RF BARNES 1075
- Mihalek RJ *Projective Geometry and Algebraic Structures* BURNETT MEYER 1072-1074
- Noether GE *Introduction to Statistics: a Fresh Approach* EL DOLNEY 335
- Nohel JA See Brauer Fred
- O'Nan Michael *Linear Algebra* DE CHRISTIE 702
- Penney DE *Perspectives in Mathematics* ET ORDMAN 568

- Ratti JS See Goodman AW
Readings for Mathematics: a Humanistic Approach ET ORDMAN 569
- Reed Michael and Simon Barry *Methods of Mathematical Physics, V. I.: Functional Analysis* JA GOLDSTEIN 1152-1153
- Reiner Irving *Introduction to Matrix Theory and Linear Algebra* DE CHRISTIE 702
- Resnikoff HL and Wells RO *Mathematics in Civilization* ET ORDMAN 568
- Roberts AW *Introductory Calculus with Analytic Geometry and Linear Algebra* 2nd ed. BURNETT MEYER 956
- Rogers Andrei *Matrix Methods in Urban and Regional Analysis* DE CHRISTIE 702
- Sandler F See Herstein IN
- Schell JF See Embry Mary
- Schneider Hans See Brauer Fred
- Semple JG See Tyrrell JA
- Simon Barry See Reed Michael
- Smirnov VI *Linear Algebra and Group Theory* DE CHRISTIE 702
- Snapper Ernst and Troyer RJ *Metric Affine Geometry* A BRUEN 330 A ADELBERG 333
- Spector Lawrence *Liberal Arts Mathematics* ET ORDMAN 569
- Spivak Michael *A Comprehensive Introduction to Differential Geometry* AD KRAMER 448
- Stammbach U See Hilton PJ
- Thomas JP See Embry Mary
- Troyer RJ See Snapper Ernst
- Tyrrell JA and Semple JG *Generalized Clifford Parallelism* A BRUEN 330
- Ward LE *Topology: An Outline for a First Course* ANN K STEHNEY 823
- Wells RO See Resnikoff HL
- Wimbish GJ *Mathematics: A Humanistic Approach* ET ORDMAN 569
- Wylie CR Jr *Introduction to Projective Geometry* BURNETT MEYER 1072-1074
- Young DM *Iterative Solution of Large Linear Systems* LA HAGEMAN 92-93
- Editorials* 821
- Editorial Notice* 475

TELEGRAPHIC REVIEWS

- Aaker David A (Ed) *Multivariate Analysis in Marketing Theory and Application* 583
- Aarhus U *Papers from the Open House for Probabilists* 228
- Aarhus U *Papers from the Open House for Functional Analysts* 461
- Achenbach JD *Contributions to the Theory of Aircraft Structures* 583
- Adams JF *Algebraic Topology--A Student's Guide* 107
- Adamson Iain T *Rings Modules and Algebras* 713
- Afifi AA Azen SP *Statistical Analysis A Computer Oriented Approach* 462
- Akivis MA Goldberg VV *Introductory Linear Algebra* 102
- Alavi Y Lick DR White AT (Ed) *Lecture Notes in Mathematics-303* 711
- Albert Arthur *Regression and the Moore-Penrose Pseudoinverse* 463
- Alder Henry L Roessler EB *Introduction to Probability and Statistics Fifth Edition* 344
- Allen Jr Richard C See Shampine LF
- Altwerger Samuel I *Modern Mathematics An Introduction* 1078
- AMS *Transactions of the Moscow Mathematical Society for the Year 1969* V 20 97
- AMS *Transactions of the Moscow Mathematical Society for the Year 1970* V 21 97
- AMS *Transactions of the Moscow Mathematical Society for the Year 1970* V 22 97
- Ambrose William G *Trigonometry A Functional Approach* 829
- Amstadter Bertram L *Reliability Mathematics Fundamentals Practices Procedures* 109
- Anderson Dan See Taylor JG
- Anderson TW Gupta SD Styan GPH A *Bibliography of Multivariate Statistical Analysis* 967
- Andree David D See Andree RV
- Andree Josephine P See Andree RV
- Andree Richard V Andree JF Andree DD *Computer Programming Techniques Analysis and Mathematics* 718
- Anger Arthur L *Computer Science The PL/1 Language* 229
- Ansorge R Tornig W *Lecture Notes in Mathematics-267* 105
- Anton Howard *Elementary Linear Algebra* 576
- Anton Howard *Elementary Linear Algebra* 338
- Applebaugh Gwendolyn Neul See Taylor JG
- Arndt Ole Jensen FV *An Introduction to a Discussion on Dialectic Materialism and Mathematics* 222
- Arnold William R See Schminke CW
- Artin Michael *Théorèmes de Représentabilité Pour Les Espaces Algébriques* 576
- Artin M Grothendieck A Verdier JL *Lecture Notes in Mathematics-269* 224
Lecture Notes in Mathematics-270 459
Lecture Notes in Mathematics-305 1081
- Artmann Benno *Eine Einführung in die Algebra* 960
- Ashley John P Harvey ER *Modern Geometry* 958
- Ashour S *Lecture Notes in Economics and Mathematical Systems-69* 461
- Athreya KB Ney PE *Branching Processes* 1086
- Aubuchon III William E See Moran Jr MM
- Aucoin Clayton V See Ohmer MM
- Auslander Louis *Mathematics Through Statistics* 574
- Averill EW *Elements of Statistics* 229
- Avila Geraldo SS *Lectures on the Wave Equation* 340
- Azen SP See Afifi AA
- Aziz AK (Ed) *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations* 341
- Babakhanian Ararat *Cohomological Methods in Group Theory* 337
- Babich VM (Ed) *Mathematical Problems in Wave Propagation Theory Part III* 113
- Bachman George See Narici L
- Bagchi TP Templeton JGC *Lecture Notes in Economics and Mathematical Systems* 72 1085
- Bailey Daniel E *Probability and Statistics Models for Research* 462
- Bailey Jr Walter L *Introductory Lectures on Automorphic Forms* 962
- Bajpai AC *Fortran and Algol A Programmed Course for Students and Technology* 837
- Bajpai OP *Foundations of Statistics* 717
- Balaam Leslie N See Federer WT
- Bar-Hillel Yehoshura See Fraenkel AA
- Barker George Phillip See Schneider H
- Barlow RE *Statistical Inference under Order Restrictions* 968

- Barnes JA See Murdoch J
- Barnett IA *Elements of Number Theory Revised Edition* 711
- Barrett James P *Elementary Computer Programs for Statistical Analysis* 582
- Battersby Albert *Network Analysis for Planning and Scheduling Third Edition* 112
- Bauer F Garabedian P Korn D *Lecture Notes in Economics and Mathematical Systems*-66 719
- Bauer Heinz *Probability Theory and Elements of Measure Theory* 108
- Bavinck H *Jacobi Series and Approximation* 579
- Baxter Willard E Sloyer Clifford W *Calculus with Probability For the Life and Management Sciences* 961
- Bear HS *Algebra and Elementary Functions* 339
- _____ *Algebra for College Students* 339
- Beard Robert M See Copi IM
- Beauchamp Murray A *Elements of Mathematical Sociology* 230
- Beauregard Raymond A Fraleigh JB A *First Course in Linear Algebra With Optional Introduction to Groups Rings and Fields* 459
- Beck Anatole (Ed) *Lecture Notes in Mathematics*-318 966
- Beckenstein Edward See Narici L
- Beckman David N See Crouch Ralph B
- Beckmann Petr *The Structure of Language A New Approach* 1088
- _____ *Orthogonal Polynomials for Engineers and Physicists* 1089
- Bedford FW Dwivedi TD *Vector Calculus* 103
- Behzad Mehdi Chartrand G *Introduction to the Theory of Graphs* 100
- Belinfante Johan GF Kolman B A *Survey of Lie Groups and Lie Algebras with Applications and Computational Methods* 102
- Belkner Horst *Metrische Räume* 1162
- Bell JL Slomson AB *Models and Ultraproducts An Introduction* 221
- Bellman Richard *Perturbation Techniques in Mathematics Physics and Engineering* 963
- _____ *Methods of Nonlinear Analysis V II* 965'
- Bendersky M *Generalized Cohomology and K-Theory* 227
- Benice Daniel D *Arithmetic and Algebra* 221
- Benjamin B Haycocks HW *The Analysis of Mortality and Other Actuarial Statistics* 719
- Benney David J See Greenspan HP
- Berberian Sterling K *Baer*-Rings* 223
- Berenstein Carlos A Dostal MA *Lecture Notes in Mathematics*-256 715
- Berman Gerald Fryer KD *Introduction to Combinatorics* 100
- Berman Simon L See Dolciani MP
- Berston Hyman Maxwell Fisher P *Collegiate Business Mathematics* 829
- Bézier P *Numerical Control Mathematics and Applications* 464
- Bhagavantam S Venkatarayudu T *Theory of Groups and Its Application to Physical Problems* 112
- Bharucha-Reid AT *Random Integral Equations* 580
- Bhat U Narayan *Elements of Applied Stochastic Processes* 108
- Bhattacharya PB Jain SK *First Course in Group Theory* 960
- Bicknell Marjorie Hoggatt Jr WE (Ed) A *Primer for the Fibonacci Numbers* 338
- Birkhoff Garrett Hall Jr M *Computers in Algebra and Number Theory* 224
- Bishop Errett Cheng H *Constructive Measure Theory* 715
- Bitter Gary G See Dorn WS
- Blackith RE Reymont RA *Multivariate Morphometrics* 464
- Blatter Jörg *Grothendieck Spaces in Approximate Theory* 341
- Blum EK *Numerical Analysis and Computation Theory and Practice* 340
- Blum Julius R Rosenblatt JI *Probability and Statistics* 462
- Bohigian Haig Edward *The Foundations and Mathematical Models of Operations Research with Extensions to the Criminal Justice System* 719
- Bolzano Bernard *Theory of Science* 457
- Boone WW Cannonito FB Lyndon RC (Ed) *Word Problems Decision Problems and the Burnside Problem in Group Theory* 834
- Borel Armand *Lecture Notes in Mathematics*-276 713
- Borsuk K *Theory of Shape* 227
- Bott R Gitler S James IM *Lecture Notes in Mathematics*-279 107
- Boullion Thomas L Odell PL *Generalized Inverse Matrices* 223
- Bourbaki Nicolas *Elements of Mathematics Commutative Algebra* 713
- Bousfield AK Kan DM *Lecture Notes in Mathematics*-304 835
- Bowen Earl K *Mathematics with Applications in Management and Economics*

- Third Edition* 583
- Bower Julia Wells *Mathematics A Creative Art* 1155
- Bowley AL *FY Edgeworth's Contributions to Mathematical Statistics* 1162
- Braverman Jerome D *Probability Logic and Management Decisions* 229
- Bredon Glen E *Introduction to Compact Transformation Groups* 343
- Breiman Leo *Statistics with a View Toward Applications* 836
- Brent Richard P *Algorithms for Minimization Without Derivatives* 341
- Bretagnolle JL *Lecture Notes in Mathematics-307* 1086
- Brewer James W Rutter EA (Ed) *Lecture Notes in Mathematics-311* 713
- Brézis H *Opérateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert* 964
- Brinker Orason L See Plachy JM
- Brodskii MS *Triangular and Jordan Representations of Linear Operators* 578
- Brody Linda A See Silverman EN
- Bronstein IN Semendyayev KA *A Guide-Book to Mathematics* 336
- Brousseau Brother Alfred *Fibonacci and Related Number Theoretic Tables* 338
- Browder William *Surgery on Simply-Connected Manifolds* 343
- Brown Sanborn C (Ed) *Changing Careers in Science and Engineering* 336
- Bruijn Nicolaas Govert de *Automath A Language for Mathematics* 838
- Brumfiel Charles F See Fleenor CR
- Bruns Carl M *Algebra An Introduction for College Students* 220
- Brunschvicg L *Les Étapes de La Philosophie Mathématique* 339
- Brush Stephen G (Ed) *Resources for the History of Physics* 832
- Brush Stephen G King AL *History in the Teaching of Physics* 832
- Bruter CP (Ed) *Lecture Notes in Mathematics-211* 100
- Bucur I *Lecture Notes in Mathematics-274* 222
- Bucur Ion Deleanu A *Introduction to the Theory of Categories and Functors* 337
- Budden FJ *The Fascination of Groups* 102
- Bunge Mario (Ed) *Problems in the Foundations of Physics* 113
- Burckel RB *Characterizations of $C(X)$ Among Its Subalgebras* 461
- Burdette AC *An Introduction to Analytic Geometry and Calculus Revised Edition* 460
- Burington Richard Stevens *Handbook of Mathematical Tables and Formulas Fifth Edition* 219
- Burke CJ See Levine G
- Burstein Samuel Z See Lax Peter D
- Burton David M *Abstract and Linear Algebra* 223
- Bush Grace A Young JE *Foundations of Mathematics Second Edition With Application to the Social and Management Sciences* 1155
- Butts Thomas *Problem Solving in Mathematics Elementary Number Theory and Arithmetic* 709
- Butzer PL Kahane J-P Szökefalvi-Nagy B *Linear Operators and Approximation* 964
- Byrne George D Hall CA (Ed) *Numerical Solution of Systems of Nonlinear Algebraic Equations* 1159
- Byron Jr Frederick W Fuller RW *Mathematics of Classical and Quantum Physics* 584
- Cain Rolene B *Elementary Statistical Concepts* 228
- Callahan John Sternberg S Weiss E *Modern Elementary Mathematics A Laboratory Approach* 221
- Cannonito FB See Boone WW
- Carico Charles C See Drooyan I
- Carter Roger W *Simple Groups of Lie Type* 459
- Cassel Don *Programming Language One* 582
- Cassels JWS *An Introduction to Diophantine Approximation* 458
- Castonguay Charles *Meaning and Existence in Mathematics* 576
- Challifour John L *Generalized Functions and Fourier Analysis An Introduction* 1160
- Chambadal L *Formulaire de Mathématiques* 709
- Chavel Isaac *Riemannian Symmetric Spaces of Rank One* 716
- Chen Wai-Kai *Applied Graph Theory* 110
- Cheng Henry See Bishop E
- Cherkasova MP *Collected Problems in Numerical Methods* 834
- Chirlian Paul M *Introduction to FORTRAN IV with Timeshare and Batch Operation* 838
- Chou Chin Cheng *Lecture Notes in Mathematics-325* 1159
- Christian Robert R *Introduction to Logic and Sets Second Edition* 1080
- Ciampi A *Classical Hamiltonian Linear Systems* 577

- Cilley David M See Kidd KP
 Cissell Helen See Cissell R
 Cissell Robert Cissell H *Mathematics of Finance Fourth Edition* 336
 Clark Colin *The Theoretical Side of Calculus* 460
 Clarke Douglas A *Foundations of Analysis with an Introduction to Logic and Set Theory* 221
 Coale Ansley J *The Growth and Structure of Human Populations A Mathematical Investigation* 582
 Cochran James Alan *The Analysis of Linear Integral Equations* 578
 Cockcroft WH *Complex Numbers A Study in Algebraic Structure* 459
 Cohen MM *A Course in Simple-Homotopy Theory* 1085
 Coles William J Reed KD Tucker DH *Calculus A Preliminary Edition* 962
 Colombo S Lavoine J *Transformations de Laplace et de Mellin* 107
 Colwell Peter Mathews JC *Introduction to Complex Variables* 1159
 Combès Michel *Fondements des Mathématiques* 99
 Comrie LJ *Chambers Shorter Six-Figure Mathematical Tables* 456
 Conover SJ *Practical Nonparametric Statistics* 581
 Constantinescu Corneliu Cornea A *Potential Theory on Harmonic Spaces* 1161
 Conte SD deBoor C *Elementary Numerical Analysis An Algorithmic Approach Second Edition* 226
 Cooley William W Lohnes PR *Multivariate Data Analysis* 581
 Copi Irving M Beard RM (Ed) *Essays on Wittgenstein's Tractatus* 832
 Corlett PN Tinsley JD *Practical Programming Second Edition* 344
 Cornea Aurel See Constantinescu C
 Cortez Marion J See Ohmer MM
 Cournot Augustin *Researches Into the Mathematical Principles of the Theory of Wealth* 465
 Cox DR *The Analysis of Binary Data* 718
 Coxeter HSM Moser WOJ *Generators and Relations for Discrete Groups Third Edition* 337
 Craggs JW *Models and Measurement* 219
 Croßley JN *What is Mathematical Logic?* 338
 Crouch Ralph B Beckman DN *The Structure of Abstract Algebra* 713
 Crowdis David G Wheeler BW *Intermediate Algebra for Colleges* 220
 Crowell Richard H See Williamson RE
 CUPM *Proceedings Summer Conference for College Teachers on Applied Mathematics* 574
 Curry Haskell B Hindley JR Seldin JP *Combinatory Logic V II* 458
 Curtain Ruth F (Ed) *Lecture Notes in Mathematics*-294 462
 Curtis Alan R *Practical Math for Business* 829
 Curtis Jr Philip C *Multivariate Calculus with Linear Algebra* 714
 ——— *Calculus With An Introduction to Vectors* 103
 Cutler Ann McShane R *The Trachtenberg Speed System of Basic Mathematics* 957
 Daclin E See Perrin J-P
 Damerau Frederick J *Markov Models and Linguistic Theory An Experimental Study of a Model for English* 583
 Dao-xing Xia *Measure and Integration Theory on Infinite-Dimensional Spaces Abstract Harmonic Analysis* 964
 Darboux Gaston *Lecons sur la Théorie Générale Des Surfaces V I-IV* 338
 Davidson Melvin PL/1 *Programming with PL/C* 837
 Davidson Ronald C Marion JB *Mathematical Preparation for General Physics with Calculus* 959
 Davies RG *Computer Programming in Quantitative Biology* 582
 Davis E Allan Pedersen JJ *Essentials of Trigonometry Second Edition* 830
 Davis Lee W *Fundamental Mathematics for Technical Students* 958
 Davis Thomas A *Algebra and Trigonometry* 97
 ——— *Algebra and Trigonometry in Four Programmed Volumes* 339
 Day A Colin *Fortran Techniques with Special Reference to Non-Numerical Applications* 838
 Day Richard H Robinson SM (Ed) *Mathematical Topics in Economic Theory and Computation* 719
 deBoor Carl See Conte SD
 deFinetti Bruno *Probability Induction and Statistics The Art of Guessing* 716
 Degrazia Joseph *Math is Fun* 710
 de la Harpe Pierre *Lecture Notes in Mathematics*-285 341
 Deleanu Aristide See Bucur I
 Dellacherie Claude *Capacités et Processus Stochastiques* 1086
 DeLuca Louis J Sedlock JT *Calculus A First Course* 1083
 Demazure Michel *Lecture Notes in*

- Mathematics-302* 712
 Denouette M See Perrin J-P
 Deskins WE *Abstract Algebra Fourth Printing* 1081
 Deuring Max *Lecture Notes in Mathematics-314* 1158
 DeVore Ronald A *Lecture Notes in Mathematics-293* 962
 Dhrymes Phoebus J *Distributed Lags Problems of Estimation and Formulation* 111
 Dick Elie M *Current Information Sources in Mathematics An Annotated Guide to Books and Periodicals 1960-1972* 574
 Dickson LE *Linear Algebras* 1081
 Dieudonné J *Elements D'Analyse Tome III* 1160
 Elements D'Analyse Tome IV 1161
 Treatise on Analysis V II 835
 Treatise on Analysis V III 1161
 Dixon Charles *Applied Mathematics of Science and Engineering* 579
 Dock V Thomas *FORTRAN IV Programming* 968
 Dodes Irving Allen *Finite Mathematics A Liberal Arts Approach* 96
 Doetsch Gustav *Guide to the Applications of the Laplace and Z-Transforms* 111
 Dolciani Mary P Berman SL Wooton W *Modern Algebra and Trigonometry Structure and Method Book Two Revised Teacher's Edition* 456
 Dolciani Mary P See Sorgenfrey RH
 Dold A *Lectures on Algebraic Topology* 342
 Dold A Eckmann B (Ed) *Lecture Notes in Mathematics-275* 225
 Lecture Notes in Mathematics-288 223
 Lecture Notes in Mathematics-317 1156
 Dorn William S Bitter GG Hector DL *Computer Applications for Calculus* 103
 Dorn William S McCracken DD *Numerical Methods with Fortran IV Case Studies* 226
 Dorsett Joseph L *College Algebra* 98
 Dorwart Harold L *The Geometry of Incidence* 342
 Dostal Milos A See Berenstein CA
 Dou Alberto *Lectures on Partial Differential Equations of First Order* 104
 Douglas Ronald G *Banach Algebra Techniques in Operator Theory* 460
 Dreyfuss Martin J *Speaking of Math Principles of Elementary Mathematics and A Study Guide* 1078
 Drooyan Irving Hadel W A *Programmed Introduction to Number Systems Second Edition* 457
 Trigonometry An Analytic Approach Second Edition 457
 Elementary Algebra Structure and Skills Third Edition 828
 Dubbey JM *Development of Modern Mathematics* 831
 Dubisch Roy See Howes VE
 DuChateau Paul *The Cauchy-Goursat Problem* 104
 Duff Charles L See Hackert AF
 Dunford Nelson Schwartz JT *Linear Operators Part III Spectral Operators* 578
 Durbin John R *Mathematics Its Spirit and Evolution* 1155
 Duren Jr William L *Calculus* 459
 Calculus and Analytic Geometry 460
 Dwinger Ph *Introduction to Boolean Algebras Second Revised and Enlarged Edition* 102
 Dwivedi TD See Bedford FW
 Dyer Eldon *Cohomology Theories* 716
 Dym H McKean HP *Fourier Series and Integrals* 1160
 Dynkin Evgenii B Yushkevich AA *Markov Processes Theorems and Problems* 967
 Eadie WT *Statistical Methods in Experimental Physics* 968
 Eames WR Stanton RG Thomas RSD (Ed) *Proceedings of the Twenty-Fifth Summer Meeting of the Canadian Mathematical Congress June 16-18 1971* 709
 Earle James H *Descriptive Geometry* 1088
 Easton Richard J Graham Jr George P *Intermediate Algebra* 339
 Eckhaus Wiktor *Matched Asymptotic Expansions and Singular Perturbations* 964
 Eckmann B See Dold A
 Edelen Dominic GB Kydonieffs AD *An Introduction to Linear Algebra for Science and Engineering* 102
 Edwards RE *Integration and Harmonic Analysis on Compact Groups* 106
 Ekambaram SK *The Statistical Basis of Quality Control Charts A Manual for Business and Factory Managers Second Revised Edition* 718
 Eley Lothar Edmund *Husserl Philosophie Der Arithmetik* 457
 Elson Mark *Concepts of Programming Languages* 968
 Elzey Freeman F A *Programmed Introduction to Statistics Second Edition* 967
 A First Reader in Statistics 229
 Emch Gerard G *Algebraic Methods in Statistical Mechanics and Quantum Field Theory* 113

- Emerson Lloyd S Paquette LR *Linear Algebra Calculus and Probability Fundamental Mathematics for the Social and Management Sciences* 710
- Emmet ER *Brain Puzzler's Delight, Fourth Edition* 710
- Enderton Herbert B A *Mathematical Introduction to Logic* 99
- Endler Otto *Valuation Theory* 458
- England AH *Complex Variable Methods in Elasticity* 465
- Erricker BC *Advanced General Statistics* 718
- Essick Edward L *RPG for System/360 and System/370* 968
- Eulenberg Milton D *Intermediate Algebra A College Approach* 221
- Even Shimon *Algorithmic Combinatorics* 833
- Everitt WN Sleeman BD (Ed) *Lecture Notes in Mathematics-280* 226
- Everling W *Lecture Notes in Economics and Mathematical Systems-65* 110
- Eves Howard W *The Other Side of the Equation* 96
- Ewald Günter *Geometry An Introduction* 227
- Eymard Pierre *Lecture Notes in Mathematics-300* 1161
- Fandel G *Lecture Notes in Economics and Mathematical Systems-76* 964
- Fang J A *Guide to the Literature of Mathematics Today* 574
- Fano Guido *Mathematical Methods of Quantum Mechanics* 230
- Farina Mario V See Gleim GA
- Farnsworth D (Ed) *Methods of Local and Global Differential Geometry in General Relativity* 720
- Federer Walter T Balaam LN *Bibliography on Experiment and Treatment Design Pre-1968* 967
- Freeman George F Brabois NR *Linear Algebra and Multivariable Calculus* 103
- Ferguson Allan (Ed) *Natural Philosophy Through the 18th Century and Allied Topics* 831
- Ferrier JP *Lecture Notes in Mathematics* 164 1085
- Fettis Henry E *An Improved Tabulation of the Plasma Dispersion Function and Its First Derivative* 1089
- Fichtenholz GM *Infinite Series Rudiments* 461
- Fierz Markus *Lecture Notes in Physics-15* 99
- Fine Terrence L *Theories of Probability An Examination of Foundations* 581
- Finlayson Bruce A *The Method of Weighted Residuals and Variational Principles with Application in Fluid Mechanics Heat and Mass Transfer* 465
- Finney DJ *An Introduction to Statistical Science in Agriculture Fourth Edition* 836
- Fisher Paul See Berston HM
- Fisher Sir Ronald A *The Design of Experiments* 1162
- Fitzgerald William M *Laboratory Manual for Elementary Mathematics Second Edition* 1080
- Flanders Harley Korfhage RR Price JJ A *First Course in Calculus with Analytic Geometry* 834
- ____ *Introductory College Mathematics with Linear Algebra and Finite Mathematics* 829
- ____ *Elementary Functions and Analytic Geometry* 1079
- Flaschel P Klingenberg W *Lecture Notes in Mathematics-282* 227
- Fleenor Charles R Shanks ME Brumfiel CF *The Elementary Functions Second Edition* 958
- Fobes Melcher P *Elementary Functions Backdrop for the Calculus* 339
- Forbes Eric G *The Unpublished Writings of Tobias Mayer V I-II* 831
- Forman William Gavurin LL *Elements of Arithmetic Algebra and Geometry* 220
- Fossum Robert M *The Divisor Class Group of a Krull Domain* 1159
- Fowler RH *The Elementary Differential Geometry of Plane Curves* 965
- Fowles Grant R *Analytical Mechanics Second Edition* 112
- Fraenkel Abraham A Bar-Hillel Y Levy A *Foundations of Set Theory Second Revised Edition* 576
- Fraïssé Roland *Cours de Logique Mathématique Tome I* 222
- Fraleigh John B *Calculus A Linear Approach V II* 103
- Fraleigh John B See Beauregard RA
- Frank Jr Charles R *Statistics and Econometrics* 717
- Frank Thomas S Smith JF *Modern Calculus* 1082
- Freedman David *Approximating Countable Markov Chains* 108
- Freiberger Walter (Ed) *Statistical Computer Performance Evaluation* 463
- Freund John E *Introduction to Probability* 836
- Friedman Avner *Differential Games* 1160
- Friend J Newton *Numbers Fun and Facts* 97

- Frisk Peter D See Gustafson RD
 Fryer KD See Berman G
 Fuchs László *Infinite Abelian Groups* V II 576
 Fuglede Bent *Lecture Notes in Mathematics* 289 715
 Fukunaga Keinosuke *Introduction to Statistical Pattern Recognition* 968
 Fuller Gordon *Analytic Geometry Fourth Edition* 829
 Fuller Robert W See Byron Jr FW
 Gagen Terrence Hale Jr MP Shult EE (Ed) *Finite Groups '72* 960
 Gallagher RH Yamada Y Oden JT (Ed) *Recent Advances in Matrix Methods of Structural Analysis and Design* 583
 Gambill Robert See Shanks ME
 Garabedian P See Bauer F
 Gardner Constance Moore See May KO
 Gardner KL Glenn JA Renton AIG (Ed) *Children Using Mathematics* 830
 Gardner Robert B *Lectures on Exterior Algebras over Commutative Rings* 224
 Garfinkel Robert S Nemhauser GL *Integer Programming* 1084
 Garnett John *Lecture Notes in Mathematics*-297 1084
 Gavurin Lester L See Forman W
 Geach PT *Logic Matters* 338
 Gear CW *Introduction to Computer Science* 838
 Gemignani Michael *Calculus A Short Course* 714
 ——— *Elementary Topology Second Edition* 966
 Geoffrion AM (Ed) *Perspectives on Optimization A Collection of Expository Articles* 579
 Germain Clarence B *PL/1 For the IBM* 360 230
 Gessner P Spremann K *Lecture Notes in Economics and Mathematical Systems*-64 106
 Chartrand Gary See Behzad M
 Giacaglia GEO *Perturbation Methods in Non-Linear Systems* 715
 Gihman II Skorohod AV *Stochastic Differential Equations* 967
 Gillespie RP *Solving Problems in Advanced Calculus I* 961
 Gilmer Robert *Multiplicative Ideal Theory* 223
 Gitler S See Bott R
 Glaeser George *Mathématiques pour l'élève professeur* 830
 Gleim George A Farina Mario V *Data Processing Mathematics* 344
 Glenn JA See Gardner KL
 Glenn William H See Johnson DA
 Glicksman Abraham M See Ruderman HD
 Gluss Brian *An Elementary Introduction to Dynamic Programming A State Equation Approach* 110
 Gobran Alfonse *Algebra A Course for College Students* 829
 Godement Roger Jacquet H *Lecture Notes in Mathematics*-260 101
 Gohring Kenneth W *Fifth Annual Simulation Symposium Progress in Simulation V 2* 1089
 Goldberg Jack L Schwartz AJ *Systems of Ordinary Differential Equations An Introduction* 105
 Goldberg VV See Akivis MA
 Goldfeld Stephen M Quandt RE *Nonlinear Methods in Econometrics* 1087
 Goldstine Herman H *The Computer From Pascal to von Neumann* 718
 Gonzalez Richard F McMillan Jr C *Machine Computation An Algorithmic Approach* 345
 Goodman AW *The Mainstream of Algebra and Trigonometry* 829
 Gordon Robert Robson JC *Krull Dimension* 1158
 Gould Henry W *Combinatorial Identities* 100
 Grabois Neil R See Feeman GF
 Graham Jr George P See Easton RJ
 Graves Robert L See Telser LG
 Grawoig Dennis See Hughes A
 Gray Andrew Lord Kelvin *An Account of His Scientific Life and Work* 831
 Gray HL Schucany WR *The Generalized Jackknife Statistic* 836
 Gray Mary W *Calculus with Finite Mathematics for Social Sciences* 225
 Greeno James G See Restle F
 Greenspan Harvey P Benney DJ *Calculus An Introduction to Applied Mathematics* 961
 Gregory Robert Todd See Young DM
 Greub Werner Halperin S Vanstone R *Connections Curvature and Cohomology V I DeRham Cohomology of Manifolds and Vector Bundles* 107 and 966
 Greyille TNE (Ed) *Population Dynamics* 582
 Griswold Ralph E *The Macro Implementation of SNOBOL 4 A Case Study of Machine-Independent Software Development* 463
 Grize Jean-Blaise *Logique Moderne Fascicule II* 99
 Grossman Michael Katz R *Non-Newtonian Calculus* 580

- Grosswald Emil See Rademacher H
 Grothendieck A See Artin M
 Gulick Denny Lipsman RL (Ed) *Lecture Notes in Mathematics*-266 106
 Gupta Somesh Das See Anderson TW
 Gustafson R David Frisk PD *Elementary Plane Geometry* 461
 Hackert Adelbert F Duff CL *Elements of Trigonometry* 97
 Hadel Walter See Drooyan I
 Hadley G Kemp MC *Finite Mathematics in Business and Economics* 220
 Hagihara Yusuke *Celestial Mechanics Perturbation Theory V II* 720
 Hajek O (Ed) *Lecture Notes in Mathematics*-235 1161
 Hájek Petr See Vořenka P
 Hakim Monique *Topos annelés et schémas relatifs* 961
 Halacy Dan Charles Babbage Father of the Computer 99
 Hale Jr Mark P See Gagen T
 Hall Charles A See Byrne GD
 Hall FM *An Introduction to Abstract Algebra V 1 Second Edition* 223
 Hall James E *Algebra A Precalculus Course* 220
 _____ *Analytic Geometry* 958
 _____ *Trigonometry Circular Functions and Their Applications* 829
 Hall Jr Marshall See Birkhoff G
 Hall Richard S *About Mathematics* 827
 Halperin Stephen See Greub W
 Hamburg Morris *Statistical Analysis for Decision Making* 228
 Hamilton Hugh J *A Primer of Complex Variables with An Introduction to Advanced Techniques Second Printing* 225
 Hansen Rodney T *Calculus It's the Limit* 103
 Happ HH (Ed) *Gabriel Kron and Systems Theory* 1087
 Harary Frank (Ed) *New Directions in the Theory of Graphs* 833
 Hardy F Lane *Essentials of Precalculus Mathematics* 959
 Hardy GH *Collected Papers of GH Hardy V V* 457
 _____ *The Integration of Functions of a Single Variable Second Edition* 834
 _____ *Orders of Infinity* 834
 Harkema R *Simultaneous Equations A Bayesian Approach* 837
 Harnett Donald L *Introduction to Statistical Methods and Solutions Manual* 109
 Harris Jr William A Sibuya Y (Ed) *Lecture Notes in Mathematics*-312 834
 Hart William L *Basic College Algebra* 98
 Hartkopf Roy *Math Without Tears Second Printing* 828
 Hartnett William E *Principles of Modern Mathematics Book 2* 340
 Harvey ER See Ashley JP
 Hashisaki Joseph See Peterson JA
 Haskell Richard E *Introduction to Vectors and Cartesian Tensors A Programmed Text for Students of Science and Engineering* 458
 Haupt Floyd E See Peterson JM
 Hawkes Nigel *The Computer Revolution* 345
 Haycocks HW See Benjamin B
 Healey James Jones M *Mathematics for Profit A Business Mathematics Text* 219
 Heaps HS *An Introduction to Computer Languages* 345
 Hector David L See Dorn WS
 Heimer Ralph T *Basic Computer Concepts A Self-Instructional Approach* 230
 Heineman E Richard *College Algebra* 574
 Heinmets F (Ed) *Concepts and Models of Biomathematics Simulation Techniques and Methods* 111
 Henderson Kenneth B Usiskin Z Zaring WM *Precalculus Mathematics* 220
 Herbrand Jacques *Jacques Herbrand Logical Writings* 832
 Hermann Armin *The Genesis of Quantum Theory (1899-1913)* 1089
 Hermann Robert *Vector Bundles in Mathematical Physics V II* 113
 Hermes Hans *Introduction to Mathematical Logic* 832
 Hershey Daniel *Transport Analysis* 1087
 Herskowitz Gerald J Schilling RB (Ed) *Semiconductor Device Modeling for Computer-Aided Design* 719
 Hesse Otto Ludwig Otto Hesse's *Gesammelte Werke* 1157
 Higgins Jon L *Mathematics Teaching and Learning* 830
 Higgins Philip J *Notes on Categories and Groupoids* 337
 Hille Einar *Methods in Classical and Functional Analysis* 578
 Hilton Peter J *Category Theory* 1081
 Hinderer K *Grundbegriffe der Wahrscheinlichkeitstheorie* 461
 Hindley JR Lercher B Seldin JP *Introduction to Combinatory Logic* 457
 Hindley J Roger See Curry HB
 Hines William W Montgomery DC *Probability and Statistics in Engineering and*

- Management Science* 836
Hironaka H Mumford D (Ed) *Oscar Zariski Collected Papers V I* 224
Hirsch Seymour C *Essentials of Fortran IV* 837
Hirst KE *Calculus of One Variable* 1083
Hocquemiller J See Weil J
Hoel Paul G Jessen RJ *Basic Statistics for Business and Economics* 344
Hoggatt Jr Werner E See Bicknell M
Hohn Franz E *Elementary Matrix Algebra Third Edition* 960
Holden Alan *Shapes Space and Symmetry* 342
Hollister Herbert A *Modern Algebra A First Course* 102
Holz Jean L See Peterson WW
Houzel C (Ed) *Lecture Notes in Mathematics-277* 226
Howes Vernon E Dubisch R *Self-Teaching Intermediate Algebra Second Edition* 828
Howson AG A *Handbook of Terms Used in Algebra and Analysis* 96
Hu TC Robinson SM (Ed) *Mathematical Programming* 964
Hu TC *Integer Programming and Network Flows* 341
Huey RM See Karbowiak AE
Hughes Ann Grawoig D *Statistics A Foundation for Analysis* 967
Hughes DR Piper FC *Projective Planes* 965
Humphreys JE *Introduction to Lie Algebras and Representation Theory* 576
Iglewicz Boris Stoyle J *An Introduction to Mathematical Reasoning* 339
Illusie Luc *Lecture Notes in Mathematics-283* 224
Ingham AE *The Distribution of Prime Numbers* 833
Ishihara Shigeru See Yano K
Iversen Birger *Lecture Notes in Mathematics-310* 1158
Jacquet Hervé *Lecture Notes in Mathematics-278* 101
Jacquet Hervé See Godement R
Jain SK See Bhattacharya PB
James IM See Bott R
Janusz Gerald J *Algebraic Number Fields* 1157
Jardine Nicholas Sibson R *Mathematical Taxonomy* 464
Jenner WE *Lectures on Non-Associative Algebras* 712
Jensen Finn V See Arndt O
Jessen Raymond J See Hoel PG
Jewett John Phelps CR *Undergraduate Education in the Mathematical Sciences 1970-71* 574
Johnson David E Johnson JR *Graph Theory with Engineering Applications* 111
Johnson Donovan A Glenn WH *Exploring Mathematics on Your Own* 709
Johnson Johnny R See Johnson DE
Johnson Norman L Kotz S *Distributions in Statistics Continuous Multivariate Distributions* 717
Johnson Phillip E A *History of Set Theory* 831
Jones Mark See Healey J
Jordan Károly *Chapters on the Classical Calculus of Probability* 108
Kahane J-P See Butzer PL
Kan DM See Bousfield AK
Kaplansky Irving *Fields and Rings Second Edition* 224
Kapur JN *The Fascinating World of Mathematics* 336
——— *Thoughts on the Nature of Mathematics* 828
——— *Thoughts on Mathematical Education* 830
Karbowiak AE Huey RM *Information Computers Machines and Man* 582
Karras U *Cutting and Pasting of Manifolds SK-Groups* 1162
Katz Robert See Grossman M
Kaufman Kenneth See Marano J
Kawata Tatsuo *Fourier Analysis in Probability Theory* 716
Keedy Marvin L Nelson CW *Geometry A Modern Introduction Second Edition* 965
Kegel Otto H Wehrfritz BAF *Locally Finite Groups* 961
Keilis-Borok VI (Ed) *Computational Seismology* 1088
Kelly GM Laplaza M Lewis G MacLane S *Lecture Notes in Mathematics-281* 224
Kemeny John G Snell JL *Mathematical Models in the Social Sciences* 230
Kemp MC See Hadley G
Kendall MG See Pearson ES
Kenyon Hewitt Morse AP *Web Derivatives* 1159
Keros John W *Computers Fortran IV and Data Processing Applications* 229
Kidd Kenneth P Myers SS Cilley DM *The Laboratory Approach to Mathematics* 576
Klendl H *Lecture Notes in Economics and Mathematical Systems-73* 964
King Allen L See Brush SG
Kingman JFC *Regenerative Phenomena* 580
Kirk Roger E (Ed) *Statistical Issues A Reader for the Behavioral Sciences* 109

- Klauder John R (Ed) *Magic Without Magic John Archibald Wheeler* 720
- Klein Erwin *Mathematical Methods in Theoretical Economics Topological and Vector Space Foundations of Equilibrium Analysis* 579
- Kleisli Heinrich *Resolutions in Additive and Non-Additive Categories* 1158
- Kline Morris *Why Johnny Can't Add The Failure of the New Math* 575
 _____ *Mathematical Thought from Ancient to Modern Times* 831
- Klingenberg W See Flaschel P
- Klinger William R Wright RR *Basic Algebra* 457
- Knops RJ (Ed) *Lecture Notes in Mathematics-316* 963
- Knutson Donald *Lecture Notes in Mathematics-308* 1081
- Kobayashi Shoshichi Nomizu K *Foundations of Differential Geometry V II* 227
 _____ *Transformation Groups in Differential Geometry* 1161
- Kochendörffer R *Determinanten und Matrizen* 459
- Kock A Wraith GC *Elementary Toposes* 227
- Kogbetliantz Ervand Krikorian A *Handbook of First Complex Prime Numbers* 223
- Kohn Joseph J *Differential Complexes* 226
- Kolchin ER *Differential Algebra and Algebraic Groups* 960
- Kolman Bernard See Belinfante JGF
- Komkov Vadim *Lecture Notes in Mathematics-253* 110
- Koosis Donald J *Business Statistics* 716
 _____ *Probability* 1086
- Korfhage Robert R See Flanders H
- Korn D See Bauer F
- Kornai János *Anti-Equilibrium On Economic Systems Theory and the Tasks of Research* 465
- Kotler Philip *Marketing Decision Making A Model Building Approach* 583
- Kotz Samuel See Johnson NL
- Kraus David H Zunde P Slamecka V *National Science Information Systems A Guide to Science Information Systems in Bulgaria Czechoslovakia Hungary Poland Romania and Yugoslavia* 97
- Krikorian Alice See Kogbetliantz E
- Krulik Stephen A *Mathematics Laboratory Handbook for Secondary Schools* 1156
 _____ *A Handbook of Aids for Teaching Junior-Senior High School Math* 1156
- Kshirsagar Anant M *Multivariate Analysis* 717
- Ku HT (Ed) *Lecture Notes in Mathematics-298*, 299 580
- Kubota Tomio *Elementary Theory of Eisenstein Series* 1157
- Kumera Antonio Spencer Donald Lie *Equations V I General Theory* 107
- Kuratowski Kazimierz *Introduction to Set Theory and Topology Second Edition* 966
- Kurepa DR (Ed) *Topology and its Applications* 343
- Kuznetsov Boris *Einstein and Dostoyevsky* 957
- Kydoniefs Anastasios D See Edelen DGB
- Lam TY *The Algebraic Theory of Quadratic Forms* 712
- Landkof NS *Foundations of Modern Potential Theory* 715
- Lane Bennie R *Programmed Guide to Accompany Finite Mathematics* 96
- Lang Serge *Introduction to Algebraic and Abelian Functions* 101
 _____ *Introduction to Algebraic Geometry* 101
- Langbehn George J Lathrop TG Martini CJ *Fundamental Concepts of Mathematics* 220
- Lanning George E See Russell DS
- Laplaza M See Kelly GM
- Larsen Ronald *An Introduction to the Theory of Multipliers* 578
- Larson Harold J *Introduction to the Theory of Statistics* 229
- Lathrop Thomas G See Langbehn GJ
- Lavoine J See Colombo S
- Lawrence J Dennis A *Catalog of Special Plane Curves* 219
- Lawson Jr Harold W See Neuhold EJ
- Lax Anneli See Lax PD
- Lax Peter D Burstein SZ Lax A *Calculus with Applications and Computing V I* 224
- Leathem JG *Volume and Surface Integrals Used in Physics* 962
- Lebedev NN *Special Functions and Their Applications* 965
- Leblanc Hugues (Ed) *Truth Syntax and Modality* 833
- Lebowitz Aaron See Rauch HE
- LeCam Lucien M Neyman J Scott EL (Ed) *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability V I-II V-VI* 344
 _____ *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability V III-IV* 463

- Leclerc Bruno *Cahiers Mathématiques IV Distributions statistiques et lois de probabilité* 967
- Ledbetter David A *Intermediate Algebra* 1156
- Ledin Jr George See Loudon RK
- Leigh Jr Egbert Giles *Adaptation and Diversity Natural History and the Mathematics of Evolution* 111
- Leithold Louis *The Calculus with Analytic Geometry Second Edition Part II* 104
- Lentin André *Équations Dans Les Monoides Libres* 1082
- Lentner Marvin *Elementary Applied Statistics* 228
- Leonard JM *Statistics The Arithmetic of Decision-Making* 1086
- Lercher B See Hindley JR
- Leslie John F Whitworth LL *Core Mathematics* 958
- L'Esperance Wilford L *Modern Statistics for Business and Economics* 228
- Levine Gustav Burke CJ *Mathematical Model Techniques for Learning Theories* 1089
- Levy Azriel See Fraenkel AA
- Lewis G See Kelly GM
- Lewis Peter AW (Ed) *Stochastic Point Processes Statistical Analysis Theory and Applications* 109
- Lial Margaret L Miller CD *College Algebra* 958
- _____ *Beginning Algebra* 828
- Lick DR See Alavi Y
- Lieberstein H Melvin *Theory of Partial Differential Equations* 340
- Lindgren Kenneth E See Wright DF
- Linsky Leonard (Ed) *Reference and Modality* 338
- Lipsman Ronald L See Gulick D
- Locke Flora M *Math Shortcuts* 220
- Loeb1 Ernest M (Ed) *Group Theory and Its Applications V II* 112
- Loeckx J *Lecture Notes in Economics and Mathematical Systems-68* 718
- Lohman Robert H *Intuitive Calculus with College Algebra* 1082
- Lohnes Paul R See Cooley WW
- Long Calvin T *Elementary Introduction to Number Theory Second Edition* 101
- Lootsma FA (Ed) *Numerical Methods for Non-Linear Optimisation* 1160
- Lorch ER *Precalculus Fundamentals of Mathematical Analysis* 575
- Lorenzen Paul *Differential and Integral A Constructive Introduction to Classical Analysis* 221
- Louden Robert K Ledin Jr G *Programming the IBM 1130 Second Edition* 463
- Lowe PG *Classical Theory of Structures Based on the Differential Equation* 111
- Lucas JR *The Concept of Probability* 108
- Luckhardt Horst *Lecture Notes in Mathematics-306* 833
- Lusin Nicolas *Les Ensembles Analytiques et Leurs Applications* 104
- Luxemburg WAJ Robinson A (Ed) *Contributions to Non-Standard Analysis* 99
- Lyapunov AA (Ed) *Systems Theory Research (Problemy Kibernetiki) V 21* 110
- Lynch Ransom V *Calculus with Computer Applications* 1082
- Lyndon RC See Boone WW
- Mackie RK *Mathematical Methods for Chemists* 1087
- MacLane S See Kelly GM
- Mahler Kurt *Introduction to p-adic Numbers and Their Functions* 711
- Mahoney Michael Sean *The Mathematical Career of Pierre de Fermat (1601-1665)* 830
- Mal'cev AI *Algorithms and Recursive Functions* 110
- _____ *The Metamathematics of Algebraic Systems Collected Papers 1936-1967* 99
- Manifold George C *Calculating with Fortran* 345
- Mann Richard A A *FORTTRAN IV Primer* 463
- Mann W Robert See Taylor Angus E
- Manougian Manoug N Northcutt RA *Ordinary Differential Equations An Introduction* 963
- Manougian MN See Ratti JS
- Mansfield Ralph *Trigonometry with Applications* 98
- Marano Joseph Kaufman K *Fundamentals of Mathematics* 828
- Marascuilo Leonard A *Statistical Methods for Behavioral Science Research* 581
- Marcus Marvin Minc H *College Algebra* 221
- _____ *Integrated Analytic Geometry and Algebra with Circular Functions* 959
- Marder L *Calculus of Several Variables* 1083
- _____ *Vector Fields* 1083
- Marion Jerry B Davidson RC *Mathematical Preparation for General Physics* 830
- Marion Jerry B See Davidson RC
- Marshall Clifford W *Applied Graph Theory* 100
- Martin BR *Statistics for Physicists* 716
- Martin Hedley G *Mathematics for Engineering Technology and Computing Science* 580

- Martini Carl J See Langbehn GJ
 Maruyama G Prokhorov YuV (Ed) *Lecture Notes in Mathematics*-330 1162
 Mather Kenneth *Statistical Analysis in Biology* 836
 Mathews Jerold C See Colwell P
 Matlis Eben *Torsion-Free Modules* 459
 ——— *Lecture Notes in Mathematics*-327 1158
 Maurin Krzysztof *Calculus of Variations and Classical Field Theory Part I* 106
 Maxfield John E Maxfield MW *Keys to Mathematics* 1078
 Maxfield Margaret W See Maxfield JE
 Maxwell Lee M Reed MB *The Theory of Graphs A Basis for Network Theory* 101
 May Francis B *Introduction to Games of Strategy* 226
 May JP *Lecture Notes in Mathematics*-271 107
 May Kenneth O *Bibliography and Research Manual of the History of Mathematics* 832
 ——— *The Mathematical Association of America Its First Fifty Years* 1157
 May Kenneth O Gardner CM (Ed) *World Directory of Historians of Mathematics First Edition* 339
 May W Graham *Linear Algebra* 712
 Mayeda Wataru *Graph Theory* 711
 McCracken Daniel D See Dorn WS
 McCullough Thomas Phillips K *Foundations of Analysis in the Complex Plane* 1084
 McElroy Elam E *Applied Business Statistics An Elementary Approach* 462
 McKean HP See Dym H
 McMillan Jr Claude See Gonzalez RF
 McNeary Samuel S *Introduction to Computational Methods for Students of Calculus* 962
 McShane Rudolph See Cutler A
 Meetham AR (Ed) *Encyclopaedia of Linguistics Information and Control* 112
 Mellor DH *The Matter of Chance* 109
 Meltzer Bernard Michie D (Ed) *Machine Intelligence* 7 582
 Mendelson Elliott *Number Systems and the Foundations of Analysis* 711
 Menges Günter *Inference and Decision* 1162
 Merriman Gaylord M Sterrett A *Matrices and Linear Systems A Programmed Introduction* 712
 Meserve Bruce E Sobel MA *Introduction to Mathematics Third Edition* 1155
 Messing William *Lecture Notes in Mathematics*-284 343
 Meyer Yves *Algebraic Numbers and Harmonic Analysis* 101
 Michie Donald See Meltzer B
 Miles John W *Integral Transforms in Applied Mathematics* 1161
 Miller Charles D See Lial ML
 Miller Ronald E *Modern Mathematical Methods for Economics and Business* 1084
 Miller Jr Willard *Symmetry Groups and Their Applications* 584
 Minc Henryk See Marcus M
 Mizrahi Abe Sullivan M *Finite Mathematics with Applications for Business and Social Sciences* 957
 Mohler RR Ruberti A (Ed) *Theory and Applications of Variable Structure Systems* 582
 Moineau J-C *Mathématique de l'esthétique* 96
 Moise Edwin E *Calculus Second Edition* 103
 ——— *Elements of Calculus Second Edition* 103
 Molk Jules See Tannery J
 Montgomery Douglas C See Hines WW
 Montias Henri *Descartes* 98
 Moon Parry *The Abacus Its history its design its possibilities in the modern world* 1155
 Moon Robert G *Applied Mathematics for Technical Programs Arithmetic and Geometry* 958
 Moore Carolyn C *Why Don't We Do Something Different?* 830
 Moore Hal G *Pre-Calculus Mathematics* 959
 Moran Jr M Marcus Aubuchon III WE *Applied Business Mathematics* 221
 Morand Max *Géométrie Spinorielle* 960
 Morgan Bryan *Men and Discoveries in Mathematics* 831
 Morse AP See Kenyon H
 Moser JK See Siegel CL
 Moser WOJ See Coxeter HSM
 Mosteller Frederick See Tanur JM
 Mulaik Stanely A *The Foundations of Factor Analysis* 229
 Mullin RC Reid KB Roselle DP (Ed) *Proceedings of the Louisiana Conference on Combinatorics Graph Theory and Computing* 100
 Mullins Jr ER Rosen D *Calculus Concepts* 1082
 Mumford D See Hironaka H
 Munem Mustafa A Tschirhart W *Intermediate Algebra* 220
 Murdoch J Barnes JA *Statistics Problems*

- and Solutions 1162
- Muroga Saburo *Threshold Logic and Its Applications* 720
- Murray W (Ed) *Numerical Methods for Unconstrained Optimisation* 1085
- Murrill Paul W Smith CL *Introduction to Computer Science* 837
 Basic Programming 230
- Myers Shirley S See Kidd KP
- Naiman Arnold Rosenfeld R Zirkel G *Understanding Statistics* 228
- Narasimhan Raghaven *Several Complex Variables* 577
- Narici Lawrence Beckenstein E Bachman G *Functional Analysis and Valuation Theory* 226
- Nayfeh Ali Hasan *Perturbation Methods* 1088
- NCTM *The Teaching of Secondary School Mathematics-33rd Yearbook* 710
 Instructional Aids in Mathematics 34th Yearbook 959
 The Slow Learner in Mathematics 35th Yearbook 457
- Nelson Charles W See Keedy ML
- Nemhauser George L See Garfinkel RS
- Neuhold Erich J Lawson Jr HW *The PL/I Machine: An Introduction to Programming* 837
- Neumann BH *Lectures on Topics in the Theory of Infinite Groups* 960
- Newton Sir Isaac *Mathematical Principles of Natural Philosophy and His System of the World V I* 98
 Philosophiae Naturalis Principia Mathematica V I-II 99
- Ney PE See Athreya KB
- Neyman Jerzy See LeCam LM
- Nijkamp P *Planning of Industrial Complexes by Means of Geometric Programming* 341
- Nilsson Nils J *Problem-Solving Methods in Artificial Intelligence* 110
- Nomizu Katsumi See Kobayashi S
- Norkin SB *Differential Equations of the Second Order with Retarded Argument* 105
- Northcutt Robert A See Manougian MN
- Novák J (Ed) *General Topology and Its Relations to Modern Analysis and Algebra IV Proceedings of the Third Prague Topological Symposium 1971* 716
- Noverraz Philippe *Pseudo-Convergence Polynomiale et Domaines d'Holomorphie en Dimension Infinite* 964
- Oberhettinger Fritz *Tables of Bessel Transforms* 580
- Odell Patrick L See Boullion TL
- Oden J Tinsley See Gallagher RH
- Ohmer Merlin M Aucoin CV Cortez MJ *Elementary Contemporary Mathematics Second Edition* 575
- OISE K-13 *Mathematics Some Non-Geometric Aspects Part II Computing Logic and Problem-Solving* 1157
- Olive Gloria *Mathematics for Liberal Arts Students* 827
- O'Neil Peter V *Fundamental Concepts of Topology* 343
- Onicescu Octav *Principes de Logique et de Philosophie Mathématique* 458
- Orey Steven *Lecture Notes on Limit Theorems for Markov Chain Transition Probabilities* 343
- Orlik Peter *Lecture Notes in Mathematics* 291 461
- Ostrowski A *Aufgabensammlung Zur Infinitesimalrechnung Band IIA Band IIB* 225
- Paley Hiram Weichsel PM *Elements of Abstract and Linear Algebra* 1081
- Panchev S *Random Functions and Turbulence* 230
- Papy *Nombres et Vectoriel Plan Reels* 341
- Paquette Laurence R See Emerson LS
- Passman Donald S *Infinite Group Rings* 102
- Patankar SV See Srinath LS
- Pearl Martin *Matrix Theory and Finite Mathematics* 712
- Pearson ES Kendall MG *Studies in the History of Statistics and Probability* 831
- Peck Lyman C *Basic Mathematics for Management and Economics* 1078
- Pedersen Jean J See Davis EA
- Pegels C Carl *BASIC A Computer Programming Language with Business and Management Applications* 838
- Pennisi Louis L *Elements of Ordinary Differential Equations* 104
- Penrose Roger *Techniques of Differential Topology in Relativity* 716
- Pepples Jr WD See Wheeler RE
- Percus JK *Combinatorial Methods* 100
- Perrin J-P Denouette M Daclin E *Switching Machines, V 1* 719
 Switching Machines, V 2 720
- Peterson John A Hashisaki J *Theory of Arithmetic Third Edition* 710
- Peterson John M Haupt FE *Intermediate Algebra and Workbook to Accompany Intermediate Algebra* 828
- Peterson W Wesley Holz JL *FORTRAN IV and the IBM 360* 581
- Petrich Mario *Introduction to Semigroups* 577

- Pfeiffer Paul E Schum DA *Introduction to Applied Probability* 835
- Phelps C Russell See Jewett J
- Phelps Jack *Elementary Mathematics Theory and Practice* 219
- Phillips EG *Functions of a Complex Variable with Applications* 963
- Phillips Keith See McCullough T
- Pietsch Albrecht *Nuclear Locally Convex Spaces Second Edition* 579
- Pinter Charles C *Set Theory* 1157
- Piper FC See Hughes DR
- Pipkin AC *Lectures on Viscoelasticity Theory* 113
- Pittnauer Franz *Lecture Notes in Mathematics-301* 963
- Pitts CGC *Introduction to Metric Spaces* 966
- Plachy Jon M Brinker OL *Elements of Algebra A Worktext Second Edition* 828
- Plotkin BI *Groups of Automorphisms of Algebraic Systems* 103
- Pokropp F *Lecture Notes in Economics and Mathematical Systems-74* 1087
- Pólya G Szegő G *Problems and Theorems in Analysis V I* 107
- Powell Alan A Williams RA (Ed) *Econometric Studies of Macro and Monetary Relations* 719
- Powers David L *Boundary Value Problems* 104
- Press S James *Applied Multivariate Analysis* 109
- Preuss Gerhard *Allgemeine Topologie* 342
- Price Justin J See Flanders H
- Proceedings *Combinatorial Mathematics and Its Applications* 100
- Prokhorov YuV See Maruyama G
- Prouse Howard L Turner VD *Principles of Mathematics* 97
- Prouse Howard L See Turner VD
- Putnam Hilary *Philosophy of Logic* 833
- Quandt Richard E See Goldfeld SM
- Quenouille MH *Rapid Statistical Calculations Second Edition* 1162
- Quine Willard Van Orman *Set Theory and Its Logic Revised Edition* 832
- Raab Joseph A *Audiovisual Materials in Mathematics* 1156
- Rademacher Hans *Topics in Analytic Number Theory* 1158
- Rademacher Hans Grosswald E *Dedekind Sums* 222
- Radó Tibor *On the Problem of Plateau Subharmonic Functions* 580
- Raghunathan MS *Discrete Subgroups of Lie Groups* 713
- Rainville Earl D *Intermediate Differential Equations Second Edition* 340
- Raj Des *The Design of Sample Surveys* 717
- Rasmussen Søren *Non-Linear Semi-Groups Evolution Equations and Productintegral Representations* 106
- Ratti JS Manougian MN *Introductory Calculus with Applications* 1082
- Rauch Harry E Lebowitz A *Elliptic Functions Theta Functions and Riemann Surfaces* 1083
- Read Ronald C A *Mathematical Background for Economists and Social Scientists* 456
- _____ *Graph Theory and Computing* 222
- Rédei L *Lacunary Polynomials Over Finite Fields* 1157
- Reed Keith D See Coles WJ
- Reed Myril B See Maxwell LM
- Reeves CM *An Introduction to Logical Design of Digital Circuits* 345
- Reid KB See Mullin RC
- Reid William H (Ed) *Mathematical Problems in the Geophysical Sciences V 1-2* 112
- Reid William T *Riccati Differential Equations* 105
- Renton AIG See Gardner KL
- Rényi Alfréd *Letters on Probability* 832
- Restle Frank Greeno JG *Introduction to Mathematical Psychology* 584
- Reyment RA See Blackith RE
- Ribenboim Paulo *Rings and Modules* 340
- Richardson Leonard F See Richardson M
- Richardson Moses Richardson LF *Fundamentals of Mathematics Fourth Edition* 575
- Richman Fred *Number Theory An Introduction to Algebra* 102
- Richman Fred Walker C Walker E *College Trigonometry* 575
- _____ *Mathematics for the Liberal Arts Student Second Edition* 1078
- Ritt Robert K *Fourier Series* 226
- Rivano Neantro Saavedra *Lecture Notes in Mathematics-265* 224
- Robert Alain *Lecture Notes in Mathematics-326* 1158
- Robertson AP Robertson W *Topological Vector Spaces Second Edition* 835
- Robertson Wendy See Robertson AP
- Robinson A See Luxemburg WAJ
- Robinson Derek JS *Finiteness Conditions and Generalized Soluble Groups* 337
- Robinson Stephen M See Day RH
- Robinson Stephen M See Hu TC
- Robison J Vincent *Modern Algebra and Trigonometry Second Edition* 456

- Robson JC See Gordon R
- Roessler Edward B See Alder HL
- Rolewicz Stefan *Metric Linear Spaces* 1161
- Rose Donald J Willoughby RA (Ed) *Sparse Matrices and Their Applications* 460
- Roselle DP See Mullin RC
- Rosen David See Mullins Jr ER
- Rosenblatt Judah I See Blum JR
- Rosenfeld Robert See Naiman A
- Ross Sheldon M *Introduction to Probability Models* 343
- Roussas George G *Contiguity of Probability Measures Some Applications in Statistics* 108
- Roxin Emilio O *Ordinary Differential Equations* 963
- Royal *Quantities Units and Symbols* 336
- Ruberti A See Mohler RR
- Rubinfoff Morris (Ed) *Advances in Computers V* 12 345
- Rubiniowicz A *Sommerfeldsche Polynom-methode* 460
- Ruderman Harry D Glicksman AM *Mathematical Systems An Introduction* 577
- Rudin Walter *Functional Analysis* 577
- Rummel RJ *Applied Factor Analysis* 584
- Russell Donald S Lanning GE *Intermediate Algebra Second Edition* 98
- Rutter Edgar A See Brewer JW
- Sagle Arthur A Walde RE *Introduction to Lie Groups and Lie Algebras* 965
- Sakai Shōichirō *C*-Algebras and W*-Algebras* 578
- Salmon Wesley C *Statistical Explanation and Statistical Relevance* 344
- Saltz Daniel A *Short Calculus An Applied Approach* 714
- Sarma KR See Srinath LS
- Sawyer WW *An Engineering Approach to Linear Algebra* 337
- Schaaf William L *The High School Mathematics Library Fifth Edition* 1156
- Schechter Martin *Principles of Functional Analysis* 106
- ____ *Spectra of Partial Differential Operators* 577
- Schey HM *Div Grad Curl and All That An Informal Text on Vector Calculus* 1083
- Schilling Ronald B See Herskowitz GJ
- Schminke CW Arnold WR (Ed) *Mathematics is a Verb Options for Teaching a Book of Readings* 1156
- Schmitt Klaus (Ed) *Delay and Functional Differential Equations and Their Applications* 340
- Schneider Hans Barker GP *Matrices and Linear Algebra Second Edition* 1081
- Schochet Claude *Cobordism From an Algebraic Point of View* 227
- Schonland David S *La Symétrie Moléculaire* 582
- Schubert Horst *Categories* 713
- Schucany WR See Gray HL
- Schum David A See Pfeiffer PE
- Schumaker John A See Weinberg GH
- Schwartz Arthur J See Goldberg JL
- Schwartz Jacob T *Introduction to Matrices and Vectors* 712
- Schwartz Jacob T See Dunford N
- Schwarz HA *Gesammelte Mathematische Abhandlungen Second Edition* 1080
- Schwarz HR *Numerical Analysis of Symmetric Matrices* 834
- Scott Elizabeth L See LeCam LM
- Searle SR *Linear Models* 717
- Sedlock James T See DeLuca LJ
- Seeley Robert T *Calculus of One Variable Second Edition* 714
- ____ *Calculus of One and Several Variables* 714
- Seip Ulrich *Lecture Notes in Mathematics-273* 341
- Seldin JF See Hindley JR
- Seldin Jonathan P See Curry HB
- Selfridge Oliver G A *Primer for FORTRAN IV On-Line* 464
- Semendyayev KA See Bronshtein IN
- Sengupta Jati K *Stochastic Programming Methods and Applications* 1160
- Sengupta Jati K See Tintner G
- Serre J-P A *Course in Arithmetic* 959
- ____ *Représentations Linéaires des Groupes Finis Deuxième édition* 224
- Sgall Petr Tesitelova M Vachek J (Ed) *Prague Studies in Mathematical Linguistics V* 3 112
- Shampine Lawrence F Allen Jr RC *Numerical Computing An Introduction* 1160
- Shanks Merrill E Gambill R *Calculus Analytic Geometry/Elementary Functions* 961
- Shanks Merrill E See Fleenor CR
- Shao Stephen P *Statistics for Business and Economics Second Edition* 462
- Sharpe DW Vámos P *Injective Modules* 337
- Shifrin Yakov Solomonovich *Statistical Antenna Theory* 1088
- Shult Ernest E See Gagen T
- Sibson Robin See Jardine N
- Sibuya Yasutaka See Harris Jr WA
- Siegel CL Moser JK *Lectures on Celestial Mechanics* 105
- Silverman Eliot N Brody LA *Statistics A Common Sense Approach* 837
- Simader Christian G *Lecture Notes in Mathematics-268* 106

- Singh Jagjit *Mathematical Ideas Their Nature and Use* 1078
- Skelton John E *An Introduction to the BASIC Language* 229
- Skorohod AV See Gihman II
- Slamecka Vladimir See Kraus DH
- Sleeman BD See Everitt WN
- Slomson AB See Bell JL
- Slook Thomas H Wurster MA *Elementary Modern Mathematics with Calculus and Computer Programming* 710
- Sloyer Clifford W See Baxter WE
- Smart James R *Modern Geometries* 1085
- Smirnov VI (Ed) *Linear Operators and Operator Equations* 227
- Smith Cecil L See Murrill PW
- Smith James F See Frank TS
- Smith Kennan T *Primer of Modern Analysis* 225
- Snell J Laurie See Kemeny JG
- Sobel Max A See Meserve BA
- Solomon Charles *Mathematics* 336
- Sorgenfrey Robert H Wooton W Dolciani MP *Modern Algebra and Trigonometry Structure and Method Book 2 New Edition Teacher's Edition* 575
- Spencer Donald See Kumpera Antonio
- Spitzbart Abraham *Analytic Geometry* 958
- Spremann K See Gessner P
- Squire William *Integration for Engineers and Scientists* 107
- Srinath LS Sarma KR Patankar SV *Basic Engineering and Mathematical Tables* 838
- Srinivasan SK Vasudevan R *Introduction to Random Differential Equations and Their Applications* 105
- Srivastava Jagdish N (Ed) *A Survey of Combinatorial Theory* 1157
- Standley Gerald B *New Methods in Symbolic Logic* 222
- Stanton RG See Eames WR
- Starke Peter H *Abstract Automata* 463
- Steen SWP *Mathematical Logic with Special Reference to the Natural Numbers* 338
- Steenrod Norman E *How to Write Mathematics* 1155
- Steger Joseph A (Ed) *Readings in Statistics for the Behavioral Scientist* 836
- Stein Sherman K *Calculus and Analytic Geometry* 1082
- Steiner Hans-Georg (Ed) *The Teaching of Geometry at the Pre-College Level* 1156
- Stenius Erik *Critical Essays* 710
- Sternberg Saul *Mathematics and Social Sciences I* 113
- Sternberg Shlomo See Callahan J
- Sterrett Andrew See Merriman GM
- Stewart GW *Introduction to Matrix Computations* 1084
- Stewart Ian *Galois Theory* 833
- Stockton Doris S *Essential Algebra* 1079
- _____ *Essential Algebra with Functions* 1079
- _____ *Essential Mathematics* 98
- Stoer Josef *Einführung in die Numerische Mathematik I* 460
- Stoyle Judith See Iglewicz B
- Strebe David D *Elements of Modern Arithmetic* 575
- Street Anne Penfold See Wallis WD
- Stroock Daniel W Varadhan SRS (Ed) *Topics in Probability Theory Seminar 1971-1972* 967
- Struble Mitch *Stretching a Point* 710
- Styan George PH See Anderson TW
- Sulanke R Wintgen P *Differentialgeometrie und Faserbündel* 966
- Sullivan Michael See Mizrahi A
- Suppes Patrick *Axiomatic Set Theory* 1080
- Sveshnikov AG Tikhonov AN *The Theory of Functions of a Complex Variable* 1084
- Swartz Clifford E *Used Math for the First Two Years of College Science* 456
- Sz-Nagy Béla (Ed) *Hilbert Space Operators and Operator Algebras* 834
- Szegö GP (Ed) *Minimisation Algorithms Mathematical Theories and Computer Results* 715
- Szegö G See Pólya G
- Szökefalvi-Nagy B See Butzer PL
- Takeuti G Zaring WM *Axiomatic Set Theory* 1080
- Tannery Jules Molk Jules *Éléments de la Théorie des Fonctions Elliptiques Second Edition Tome I-IV*
- Tanur Judith M Mosteller F (Ed) *Statistics A Guide to the Unknown* 109
- Tarski Alfred *Introduction a la Logique* 222
- Tatsuoka Maurice M *Multivariate Analysis Techniques for Educational and Psychological Research* 462
- Taylor Angus E Mann WR *Advanced Calculus Second Edition* 225
- Taylor Howard E Wade TL *Contemporary Trigonometry* 829
- Taylor Joan Gary Applebaugh GN Anderson D *Finite Mathematics* 827
- Taylor Joseph L *Measure Algebras* 1159
- Teague Robert *Computing Problems for Fortran Solution* 463
- Teekens R *Prediction Methods in*

- Multiplicative Models* 344
- Teixeira F Gomes *Traité des courbes spéciales remarquables planes et gauches Tome I-III* 342
- Telling HG *The Rational Quartic Curve in Space of Three and Four Dimensions* 966
- Telser Lester G Graves RL *Functional Analysis in Mathematical Economics Optimization Over Infinite Horizons* 1087
- Templeton JGC See Bagchi TP
- Tesitelová Marie See Sgall P
- Theil Henri *Principles of Econometrics* 464
- *Statistical Decomposition Analysis* 584
- Thomas Ann M See Thomas JW
- Thomas Jr George B *Calculus and Analytic Geometry Alternate Edition* 714
- Thomas James W Thomas AM *Finite Mathematics* 957
- Thomas John B *An Introduction to Applied Probability and Random Processes* 228
- Thomas RSD See Eames WR
- Thompson Colin J *Mathematical Statistical Mechanics* 112
- Thompson Gerald E *Linear Programming An Elementary Introduction* 226
- Thompson Howard E *Applications of Calculus in Business and Economics* 1087
- Thurston Hugh *The Calculus An Introduction* 714
- Tijms HC *Analysis of (s,S) Inventory Models* 1087
- Tikhonov AN See Sveshnikov AG
- Tinsley JD See Corlett PN
- Tintner Gerhard Sengupta JK *Stochastic Economics Stochastic Processes Control and Programming* 583
- Topping David M *Lectures on Von Neumann Algebras* 964
- Törnig W See Ansorge R
- Tou Julius T (Ed) *Advances in Information Systems Science V 4* 110
- Tougeron Jean Claude *Idéaux de fonctions différentiables* 341
- Tranter CJ *Integral Transforms in Mathematical Physics* 583
- Treiman Sam B *Lectures on Current Algebra and Its Applications* 1089
- Triola Mario F *Mathematics and the Modern World* 1079
- Trotter Hale F See Williamson RE
- Tschebyscheff PL *Theorie der Congruenzen (Elemente der Zahlentheorie)* 338
- Tschirhart William See Munem MA
- Tucker Don H See Coles WJ
- Turner JS *Buoyancy Effects in Fluids* 720
- Turner V Dean Prouse HL *Introduction to Mathematics* 98
- Turner V Dean See Prouse HL
- Ullman Neil R *Statistics An Applied Approach* 228
- US Army *Transactions of the Sixteenth Conference of Army Mathematicians* 97
- *Transactions of the Seventeenth Conference of Army Mathematicians* 709
- Usiskin Zalman See Henderson KB
- Vachek Josef See Sgall P
- Vámos P See Sharpe DW
- van der Merwe Alwyn See Yourgrau W
- Van Note Peter *Tangrams Picture-Making Puzzle Game* 709
- Vanstone Ray See Greub W
- Varadhan SRS See Stroock DW
- Varaiya PP *Notes on Optimization* 835
- Vasudevan R See Srinivasan SK
- Venkatarayudu T See Bhagavantam S
- Verbeek A *Superextensions of Topological Spaces* 1085
- Verdier JL See Artin M
- Vervaat W *Success Epochs in Bernoulli Trials With Applications in Number Theory* 1086
- Vick James W *Homology Theory An Introduction to Algebraic Topology* 835
- Vilenkin NYa *Functional Analysis* 579
- Voßenka Petr Hájek P *The Theory of Semisets* 222
- Wade Thomas L See Taylor HE
- Waelbroeck Lucien (Ed) *Lecture Notes in Mathematics-331* 1160
- Walde Ralph E See Sagle AA
- Walker Carol See Richman F
- Walker Elbert See Richman F
- Wallace Philip R *Mathematical Analysis of Physical Problems* 1089
- Wallis Jennifer Seberry See Wallis WD
- Wallis Jennifer Wallis WD (Ed) *Proceedings of the First Australian Conference on Combinatorial Mathematics* 458
- Wallis WD Street AP Wallis JS *Lecture Notes in Mathematics-292* 711
- Wallis WD See Wallis J
- Walsh T *On Summability Methods for Conjugate Fourier-Stieltjes Integrals in Several Variables and Generalizations* 1159
- Walter Wolfgang *Gewöhnliche Differentialgleichungen* 963
- Ward Jr Lewis E *Topology An Outline for a First Course* 343
- Wardle ME *Computing in Mathematics From Problem to Program* 345

- Warner Garth *Harmonic Analysis on Semi-Simple Lie Groups I* 105
 Harmonic Analysis on Semi-Simple Lie Groups II 577
- Wasan MT *Mathematical Probability* 581
- Washington Allyn J *An Introduction to Calculus with Applications* 225
 Mathematics A Developmental Approach 339
- Watanabe Satoshi (Ed) *Frontiers of Pattern Recognition* 838
- Watson GN *Complex Integration and Cauchy's Theorem* 1084
- Webber G Cuthbert *Algebraic Structures for Teachers* 959
- Wehrfritz Bertram AF See Kegel OH
- Weichsel Paul M See Paley H
- Weil J Hocquemiller J *Algèbre Solutions Développées des exercices* 712
- Weinberg George H Schumaker JA *Statistics An Intuitive Approach Second Edition* 228
- Weinberg Gerald M *The Psychology of Computer Programming* 718
- Weinberger HF A *First Course in Partial Differential Equations With Complex Variables and Transform Methods* 963
- Weinstein Raoul L *Precalculus Mathematics A Fundamental Approach* 959
- Weiss Edwin See Callahan J
- Wells Jr RO *Differential Analysis on Complex Manifolds* 962
- Welsh DJA Woodall DR (Ed) *Combinatorics* 711
- Wheeler Brandon W See Crowdis DG
- Wheeler Ruric E Peeples Jr WD *Modern Mathematics for Business Students* 957
- Wheeler Ruric E *Fundamental College Mathematics Number Systems and Intuitive Geometry* 1080
 Modern Mathematics An Elementary Approach Third Edition 957
- White AT See Alavi Y
- Whitehead AN *The Axioms of Descriptive Geometry* 965
- Whitehead Jr Earl Glen *Enumerative Combinatorics 1971-1972* 100
 Combinatorial Algorithms 833
- Whiteside DT (Ed) *The Mathematical Papers of Isaac Newton V 5* 98
- Whitesitt J Eldon *Principles of Modern Algebra Second Edition* 960
- Whitney Hassler *Complex Analytic Varieties* 1085
- Whitworth Larry L See Leslie JF
- Wigner Eugene P *Symmetries and Reflections Scientific Essays of Eugene P Wigner* 827
- Wilde Daniel U *An Introduction to Computing Problem-Solving Algorithms and Data Structures* 838
- Wilenskin NJ *Unterhaltsame Mengenlehre* 1080
- Willerding Margaret F *A First Course in College Mathematics and A First Course in College Mathematics A Work-text* 958
- Williams CB *Style and Vocabulary Numerical Studies* 1088
- Williams Gareth A *A Course in Linear Algebra* 833
- Williams IP *Matrices for Scientists* 712
- Williams J *Complex Numbers* 1083
- Williams K *Problems in Statistics The Poisson and Exponential Distributions* 344
- Williams Keith W *Introduction to College Mathematics* 828
- Williams Ralph C *Mathematics for Communication Number Relations* 1079
- Williams Ross A See Powell AA
- Williamson Richard E Crowell RH Trotter HF *Calculus of Vector Functions Third Edition* 104
- Willoughby Ralph A See Rose DJ
- Winkloughby Stephen S *Statistics and Probability* 836
- Wilson Robin J *Introduction to Graph Theory* 711
- Wintgen P See Sulanke R
- Wisner Robert J *Elements of Probability* 1086
- Witter George E *The Structure of Mathematics An Introduction* 219
- Wolf Joseph A *Spaces of Constant Curvature Second Edition* 342
- Wolfe Carvel S *Linear Programming with Fortran* 835
- Wolff Peter *Breakthroughs in Mathematics* 827
- Woodall DR See Welsh DJA
- Wooton William See Sorgenfrey RH
- Wooton William See Dolciani MP
- Wooton William *Modern Trigonometry Revised Edition* 575
- Wraith GC See Kock A
- Wrede Robert C *Introduction to Vector and Tensor Analysis* 1083
- Wren F Lynwood *Basic Mathematical Concepts Second Edition* 1156
- Wright D Franklin Lindgren KE *Intermediate Algebra for College Students* 97
 Elementary Algebra for College Students 97

- Wright Robert R See Klinger WR
 Wurster Marie A See Slook TH
 Yaglom IM *Geometric Transformations*
 III 580
 Yamada Y See Gallagher RH
 Yano Kentaro Ishihara S *Tangent and Co-
 tangent Bundles Differential Geome-
 try* 1161
 Yeh Rui Zong *Modern Probability Theory*
 1086
 Young David M Gregory RT *A Survey of
 Numerical Mathematics 2 Vols* 1159
 Young Grace Chisholm See Young WH
 Young John E See Bush GA
 Young WH Young GC *The Theory of Sets
 of Points Second Edition* 576
 Young WH *The Fundamental Theorems of
 the Differential Calculus* 1085
 Yourgrau Wolfgang van der Merwe A (Ed)
 *Perspectives in Quantum Theory Es-
 says in Honor of Alfred Landé* 1088
 Yushkevich Aleksandr A See Dynkin EB
 Zacks Shelemyahu *The Theory of Statis-
 tical Inference* 581
 Zagier Don Bernard *Lecture Notes in
 Mathematics-290* 461
 Zarantonello Eduardo H (Ed) *Contribu-
 tions to Nonlinear Functional Analy-
 sis* 579
 Zaremba SK (Ed) *Applications of Number
 Theory to Numerical Analysis* 105
 Zarembka Paul *Toward a Theory of Econo-
 mic Development* 464
 Zaring Wilson M See Henderson KB
 Zaring WM See Takeuti G
 Zelazko W *Selected Topics in Topologi-
 cal Algebras* 106
 Banach Algebras 835
 Zelinsky Daniel A *First Course in Li-
 near Algebra Second Edition* 459
 Zellner Arnold *An Introduction to Baye-
 sian Inference in Econometrics* 465
 Zemanian AH *Realizability Theory for
 Continuous Linear Systems* 461
 Zierer Ernesto *The Theory of Graphs in
 Linguistics* 112
 Zimmer Horst G *Lecture Notes in Mathe-
 matics-262* 101
 Zirkel Gene See Naiman A
 Zlot William *Sourcebook of Fundamental
 Mathematics Series Arithmetic* 1079
 *Sourcebook of Fundamental Mathe-
 matics Series Elementary Algebra*
 1079
 *Sourcebook of Fundamental Mathe-
 matics Series Elementary Geometry*
 1079
 Zubrzycki Stefan *Lectures in Probabil-
 ity Theory and Mathematical Statis-
 tics* 717
 Zunde Pranas See Kraus DH

NEWS AND NOTICES

PERSONAL ITEMS

114, 231, 346, 466, 585, 721, 839, 969, 1090, 1163

GENERAL INFORMATION

ACM George E. Forsythe Student Paper Competition 585	Mathematical Research and Education 586
All-College Conference Room to honor David W. Blakeslee 346	Fourteenth Biennial International Seminar of the Canadian Mathematical Congress 346
Conference on the Application of Undergraduate Mathematics in the Life, Managerial, Social and Engineering Sciences 467	New Sabbatical Leave Exchange Service 840
Conference on the Influence of Computing on	Seminar on generalized inverses and applications at the MRC, University of Wisconsin 840
	USA Mathematical Olympiad 231
	Unsolved problems in mathematics 841

NECROLOGY

Ayer Miriam C 969	Levy BR 1163
Carter HC 969	Loring RJ 1090
Curtis HB Jr. 969	Macdonald SL 1163
Daus PH 1163	Moursund AF 467
Dean AE 969	Price HV 1163
Earl JM 467	Snyder AD 1090
Edwards PD 1163	Vandiver HS 1163
Foster JF 1163	Watt MW 1163
Graves LM 1163	Winger RM 1163
Lefschetz Solomon 467	

**REPORTS AND ANNOUNCEMENTS OF THE ASSOCIATION AND ITS SECTIONS
MEETINGS AND ANNOUNCEMENTS OF THE ASSOCIATION**

- Academic members elected into the Association
HL ALDER 597
- Acknowledgment 1184-1185
- Announcement of Lester R. Ford Awards
HL ALDER 849
- Announcement of W. B. Ford Lecture Fund 595
- Committee on Educational Media 347
- Disability Income Plan added to the MAA
Group Insurance Program 844
- Employment Information for Mathematicians
115 1090
- Fifty-fourth Summer Meeting of the Association
HL ALDER 1164-1179
- Fifty-sixth Annual Meeting of the Association
HL ALDER 587-597
- Films produced by the MAA 849
- Honorary Life Membership for Professor
EP Starke JOSHUA BARLAZ 1181-1182
- MAA publishes Guidelines for evaluating
college mathematics programs 841
- Mathematical Sciences Employment Register —
Open Register 1183
- New Sectional Governors of the Association
AB WILLCOX 848
- Officers and Committees as of February 1, 1973
468-473
- Proceedings of the 1971 Summer Conference
held at the University of Missouri, Rolla 347
- The Putnam Mathematical Competition 849
- Report of the Treasurer for the year 1972
LEONARD GILLMAN 1180

MEETINGS OF ITS SECTIONS

- Allegheny Mountain May 1973 MR WOODARD
1094
- Florida March 1973 FL CLEAVER 847
- Illinois May 1973 H SAAR 975
- Indiana May 1972 RT HOOD 114 November
1972 RT HOOD 844 April 1973 RT HOOD 1091
- Iowa April 1973 BE GILLAM 972
- Louisiana-Mississippi February 1973 PL FORD
845
- Maryland-District of Columbia-Virginia Novem-
ber 1972 JM SMITH 723 April 1973 JM
SMITH 1182
- Metropolitan New York April 1973 RORA
IACOBACCI 1092
- Nebraska April 1973 HM COX 972
- North Central October 1972 HM ANDERSON
467 April 1973 HM ANDERSON 973
- Northeastern June 1973 GW BEST 1183
- Northern California February 1972 NEWMAN
FISHER 722 February 1973 NEWMAN FISHER
846
- Ohio November 1972 RH ROLWING 586 April
1973 RH ROLWING 973
- Oklahoma-Arkansas April 1973 EK McLACH-
LAN 974
- Philadelphia November 1972 AE FILANO 723
- Rocky Mountain May 1973 DJ STERLING 1096
- Seaway November 1972 EMMET STOPHER 587
May 1973 EMMET STOPHER 1097
- Southeastern March 1973 JD NEFF 969
- Southern California March 1973 TN ROBERTSON
971
- Southwestern April 1973 A SWIMMER 1093
- Texas April 1973 JC BRADFORD 1093